Closed-Loop Control for Cardiopulmonary Management and Intensive Care Unit Sedation using Digital Imaging

A Dissertation Presented to
The Academic Faculty of
The School of Aerospace Engineering

by

Behnood Gholami

In Partial Fulfillment of
The Requirements for the Degree of
Doctor of Philosophy in Aerospace Engineering

Georgia Institute of Technology
August 2010

Closed-Loop Control for Cardiopulmonary Management and Intensive Care Unit Sedation using Digital Imaging

Approved by:

Dr. Wassim M. Haddad, Chairman
Aerospace Engineering
Georgia Institute of Technology

Dr. Allen R. Tannenbaum
Electrical & Computer and
Biomedical Engineering
Georgia Institute of Technology

Dr. James M. Bailey
Department of Anesthesiology
Northeast Georgia Medical Center

Dr. Eric Feron
Aerospace Engineering
Georgia Institute of Technology

Dr. John-Paul Clarke
Aerospace Engineering
Georgia Institute of Technology

Date Approved: June 25, 2010

To my parents Abolfazl Gholami and Mahboubeh Nadjafkhani, and my wife Negar Tavassolian

Up from Earth's Centre through the seventh Gate
I rose, and on the Throne of Saturn sate
And many Knots unravel'd by the Road
But not the Knot of Human Death and Fate

Omar Khayyam (1048–1131 AD, Neyshapur, Iran)
Persian mathematician, philosopher, and poet
(Translation by E. FitzGerald)

Acknowledgements

I would like to take this opportunity and thank several people without whom this work would not be possible.

First and foremost, I would like to sincerely thank my advisor and mentor Professor Wassim M. Haddad. He is an example of an individual who appreciates science for the sake of science, and pursues perfection in all aspects of life. His personality, philosophy of life, and pursuit of scholarly perfection has truly inspired me. I am grateful for his unconditional support and guidance throughout the period of my doctoral work at Georgia Tech. I also thank him for encouraging rigor, strengthening my love of mathematics, and for giving me the freedom to pursue my academic interests. I thank him for his friendship, his insights, and his help in allowing me to identify my strengths and minimize my weaknesses. As he has always said, education is a privilege, and I truly had the privilege of being his student. I also thank his wife Lydia Haddad, not only for showing interest in my research and alerting me to the infant COPE database, but also for her genuine friendship.

I also would like to express my gratitude to my co-advisor, Professor Allen R. Tannenbaum. His modest personality not only is an inspiration to me, but truly reflects his achievements in science and life in general. He has a rare combination of extraordinary insight and analytical capability while having a welcoming demeanor. I am sincerely indebted to him for believing in me, investing in me, and encouraging me to grow and explore new ideas. It was truly an honor being his student during this

period. I also would like to thank him for helping me cross the bridge from control theory to the world of computer vision and medical image processing.

I am deeply grateful to Dr. James M. Bailey. His clinical expertise was invaluable to this work. The work in this dissertation would by no means be possible without his ideas, insights, and suggestions. I thank him for giving me the opportunity to meet with him regularly and reviewing my work given his overwhelming time constraints.

I am thankful to Professors Eric Feron and John-Paul Clarke for taking time to serve on my dissertation reading committee and for their useful comments and suggestions to improve this dissertation. I would like to thank Professor Sheryl Brahnam for providing the infant COPE database.

I am grateful to my brother Farnood Gholami and sister Farnaz Gholami whose passion to pursue their dreams has always been an example for me.

I am deeply grateful to my wife and my soul mate Negar Tavassolian. I admire her exemplary modesty and effort to achieve perfection. I am indebted for her love and support throughout these years that made my stay at Georgia Tech extraordinary. This work would not be possible without her.

Lastly, but by no means least, I would like to express my deepest gratitude to my parents and best friends Abolfazl Gholami and Mahboubeh Nadjafkhani, for their unconditional love, support, and encouragement, and for instilling in me the passion for learning. They have been and continue to be examples of people who passionately pursued their dreams. They are my role-models; individuals who have "unraveled the knot" in the purpose of life. They are a constant source of inspiration to me.

Table of Contents

List of Tables

List of Figures

Summary

This dissertation introduces a new problem in the delivery of healthcare, which could result in lower cost and a higher quality of medical care as compared to the current healthcare practice. In particular, a framework is developed for sedation and cardiopulmonary management for patients in the intensive care unit. A method is introduced to automatically detect pain and agitation in nonverbal patients, specifically in sedated patients in the intensive care unit, using their facial expressions. Furthermore, deterministic as well as probabilistic expert systems are developed to suggest the appropriate drug dose based on patient sedation level. This framework can be used to automatically control the level of sedation in the intensive care unit patients via a closed-loop control system. Specifically, video and other physiological variables of a patient can be constantly monitored by a computer and used as a feedback signal in a closed-loop control architecture. In addition, the expert system selects the appropriate drug dose based on the patient's sedation level.

Patients in the intensive care unit who require mechanical ventilation due to acute respiratory failure also frequently require the administration of sedative agents. The need for sedation arises both from patient anxiety due to the loss of personal control and the unfamiliar and intrusive environment of the intensive care unit, and also due to pain or other variants of noxious stimuli. While physicians select the agent(s) used for sedation and cardiovascular function, the actual administration of these agents is the responsibility of the nursing staff. If clinical decision support systems and closed-

loop control systems could be developed for critical care monitoring and lifesaving interventions as well as the administration of sedation and cardiopulmonary management, the intensive care unit nurse could be released from the intense monitoring of sedation, allowing her/him to focus on other critical tasks. One particularly attractive strategy is to utilize the knowledge and experience of skilled clinicians, capturing explicitly the rules expert clinicians use to decide on how to titrate drug doses depending on the level of sedation. In this dissertation, we develop a rule-based expert system for cardiopulmonary management and intensive care unit sedation. Furthermore, we use probability theory to quantify uncertainty and to extend the proposed rule-based expert system to deal with more realistic situations.

Pain assessment in patients who are unable to verbally communicate is a challenging problem. The fundamental limitations in pain assessment stem from subjective assessment criteria, rather than quantifiable, measurable data. This often results in poor quality and inconsistent treatment of patient pain management. Recent advancements in pattern recognition techniques using relevance vector machine learning techniques can assist medical staff in assessing pain by constantly monitoring the patient and providing the clinician with quantifiable data for pain management. The relevance vector machine (RVM) classification technique is a Bayesian extension of the support vector machine (SVM) algorithm which achieves comparable performance to SVM while providing posterior probabilities for class memberships and a sparser model. If classes represent "pure" facial expressions (i.e., extreme expressions that an observer can identify with a high degree of confidence), then the posterior probability of the membership of some intermediate facial expression to a class can provide an estimate of the intensity of such an expression. In this dissertation, we use the RVM classification technique to distinguish pain from non-pain as well as assess pain intensity levels. We also correlate our results with the pain intensity assessed by expert

and non-expert human examiners.

Next, we consider facial expression recognition using an unsupervised learning framework. Specifically, given a data set composed of a number of facial images of the same subject with different facial expressions, the algorithm segments the data set into groups corresponding to different facial expressions. Each facial image can be regarded as a point in a high-dimensional space, and the collection of images of the same subject resides on a manifold within this space. We show that different facial expressions reside on distinct subspaces if the manifold is unfolded. In particular, semi-definite embedding is used to reduce the dimensionality and unfold the manifold of facial images. Next, generalized principal component analysis is used to fit a series of subspaces to the data points and associate each data point to a subspace. Data points that belong to the same subspace are shown to belong to the same facial expression.

In clinical intensive care unit practice sedative/analgesic agents are titrated to achieve a specific level of sedation. The level of sedation is currently based on clinical scoring systems. Examples include the motor activity assessment scale (MAAS), the Richmond agitation-sedation scale (RASS), and the modified Ramsay sedation scale (MRSS). In general, the goal of the clinician is to find the drug dose that maintains the patient at a sedation score corresponding to a moderately sedated state. This is typically done empirically, administering a drug dose that usually is in the effective range for most patients, observing the patient's response, and then adjusting the dose accordingly. However, the response of patients to any drug dose is a reflection of the pharmacokinetic and pharmacodynamic properties of the drug and the specific patient. In this research, we use pharmacokinetic and pharmacodynamic modeling to find an optimal drug dosing control policy to drive the patient to a desired MRSS score.

Linear matrix inequalities provide a powerful design framework for linear control problems. The dynamical models of many biological, pharmacological, and physiological processes such as pharmacokinetics [7, 142], metabolic systems [27], epidemic dynamics [83, 84], biochemical reactions [34, 95], endocrine systems [27], and lipoprotein kinetics [95] are derived from mass and energy balance considerations that involve dynamic states whose values are nonnegative. Hence, it follows from physical considerations that the state trajectory of such systems remains in the nonnegative orthant of the state space for nonnegative initial conditions. Such systems are commonly referred to as *nonnegative dynamical systems*. In this part of the dissertation, we use linear matrix inequalities to develop $\mathcal{H}_2$ (sub)optimal estimators and controllers for nonnegative dynamical systems. Specifically, we formulate a series of generalized eigenvalue problems subject to a set of linear matrix inequality constraints for designing $\mathcal{H}_2$ suboptimal estimators, static controllers, and dynamic controllers for nonnegative dynamical systems. The resulting $\mathcal{H}_2$ suboptimal controllers guarantee that the closed-loop plant system states remain in the nonnegative orthant of the state space.

Atrial fibrillation, a cardiac arrhythmia characterized by unsynchronized electrical activity in the atrial chambers of the heart, is a rapidly growing problem in modern societies. One treatment, referred to as *catheter ablation*, targets specific parts of the left atrium for radio frequency ablation using an intracardiac catheter. Magnetic resonance imaging has been used for both pre- and and post-ablation assessment of the atrial wall. Magnetic resonance imaging can aid in selecting the right candidate for the ablation procedure and assessing post-ablation scar formations. Image processing techniques can be used for automatic segmentation of the atrial wall, which facilitates an accurate statistical assessment of the region. As a first step towards the general solution to the computer-assisted segmentation of the left atrial wall, we use shape

learning and shape-based image segmentation to identify the endocardial wall of the left atrium in the delayed-enhancement magnetic resonance images.

Chapter 1

Introduction

Modern control technology is having a revolutionary impact in modern medicine through medical robotics (stereotactical brain surgery, implant fitting, and coronary procedures), electrophysiological systems (pacemakers and automatic implantable defibrillators), life support (ventilators and artificial hearts), and medical imaging (image-guided surgery and therapy). An additional area of medicine that can benefit enormously from systems and control oriented ideas is clinical pharmacology, in which mathematical modeling plays a prominent role [3, 34, 58, 60, 61, 82, 123]. This is particularly true when dealing with critically ill patients in the intensive care unit (ICU) or in the operating room. These patients often require administration of drugs to regulate key physiological variables, such as level of consciousness, heart rate, blood pressure, ventilatory drive, etc., within desired targets. The rate of administration of these drugs is critical, requiring constant monitoring and frequent adjustments. Open-loop control by clinical personnel can be tedious, imprecise, time-consuming, and sometimes of poor quality. Hence, the need for closed-loop control (active control) in medical drug delivery systems is significant, with the potential for improving the quality of medical care as well as curtailing the increasing cost of health care.

One of the main drawbacks in developing active control-based drug delivery systems is the lack of accurate mathematical models for characterizing the dynamic

behavior of drugs on physiological variables. System nonlinearities, model parameter variations from patient to patient, as well as parameter variations within the same patient under different conditions make it very challenging to develop models and effective control law architectures for active drug delivery systems. Standard data-driven system identification techniques may not be applicable to complex biological system modeling involving *in situ* diagnostics.

Patients in the intensive care unit who require mechanical ventilation due to acute respiratory failure also frequently require the administration of sedative agents. The need for sedation arises from patient anxiety due to the loss of personal control and the unfamiliar and intrusive environment of the intensive care unit. In addition, pain or other variants of noxious stimuli frequently require administration of anxiolytic and analgesic drugs for patient comfort. In particular, the interface between the patient and the ventilator is typically an endotracheal tube passing through the oropharynx and into the trachea. Due to the powerful gag reflex, this tube is very noxious. Without sedation patients can become dangerously agitated, risking dislodgement of life support devices in the worst case and, in any case, resulting in stress that is ethically unacceptable and also physiologically unacceptable due to deleterious increases in heart rate, blood pressure, and work of breathing.

The current clinical standard in the ICU for assessing the level of sedation in adults is an ordinal scoring system, such as the motor activity and assessment scale (MAAS) [40] or the Richmond agitation-sedation scale (RASS) [129], which includes the assessment of the level of agitation of the patient as well as the level of consciousness. For example, the MAAS system evaluates the level of sedation and agitation on a score of 0-6 as follows: 0 - unresponsive; 1 - responsive only to noxious stimuli; 2 - responsive to touch or name; 3 - calm and cooperative; 4 -restless and cooperative; 5 - agitated; and 6 - dangerously agitated.

Assessment of the level of sedation and agitation of a patient is, therefore, subjective and limited in accuracy and resolution, and hence, prone to error in assessing the level of sedation which in turn may lead to oversedation. Oversedation increases risk to the patient since liberation from mechanical ventilation, one of the most common life-saving procedures performed in the ICU, may not be possible due to a diminished level of consciousness and respiratory depression from sedative drugs resulting in prolonged length of stay in the ICU. Prolonged ventilation is expensive and is associated with known risks, such as inadvertent extubation, laryngo-tracheal trauma, and ventilator-associated pneumonia. Alternatively, undersedation leads to agitation and can result in dangerous situations for both the patient and the intensivist. Specifically, agitated patients can do physical harm to themselves by dislodging their endotracheal tube which can potentially endanger their life.

Although a number of algorithms have been developed and implemented for the problem of closed-loop control of general anesthesia [66, 71, 87, 124, 133, 147, 150], the problem of closed-loop control of ICU sedation is undeveloped [60]. This is mainly due to the inherent challenges in the problem of ICU sedation, where there is considerable uncertainty in the environment (as opposed to the controlled environment of an operating room) and the treatments are primarily based on subjective assessments. While the bispectral index (BIS)—a derivatitive of the electroencephalogram signal—is used for assessing the level of conciousness in general anesthesia, no single objective measurement of sedation has been developed. In addition, several studies have shown inconsistent results for using the BIS score for sedation assessment [72, 111, 112, 122, 131, 132]. As a result, closing the loop in the ICU involves developing a control architecture to reliably regulate a patient's sedation level around a desired state as well as designing patient monitoring systems to assess the level of sedation and agitation, and provide the controller with a viable feedback signal.

In this dissertation, we address problems in controller design as well as sedation and agitation assessment for cardiopulmonary management and ICU sedation. In particular, deterministic as well as probabilistic expert systems are developed to suggest the appropriate drug dose based on patients sedation level. This framework can be used to automatically control the level of sedation in intensive care unit patients via a closed-loop control system. Specifically, video and other physiological variables of a patient can be constantly monitored by a computer and used as a feedback signal in a closed-loop control architecture. In addition, the framework is also applicable to clinical decision support systems for cardiopulmonary management and ICU sedation.

Computer vision techniques offer the possibility to quantify agitation in sedated ICU patients. In particular, such techniques can be used to develop objective agitation measurements from patient motion. In the case of paraplegic patients, whole body movement is not available, and hence, monitoring the whole body motion is not a viable solution. In this case, measuring head motion and facial grimacing for quantifying patient agitation and sedation in critical care can be a useful alternative. Of course, patient occlusions due to medical equipment will need to be accounted for within the machine learning algorithms.

Machine learning techniques can potentially be useful in assessing sedation and agitation in the ICU. The fundamental limitations in sedation and agitation assessment in the ICU stem from subjective assessment criteria, rather than quantifiable, measurable data for ICU sedation. This often results in poor quality and inconsistent treatment of patient agitation. Advances in computer vision techniques can potentially assist the medical staff in assessing sedation and agitation by constantly monitoring the patient and providing the clinician with quantifiable data for ICU sedation. An automatic sedation and pain assessment system can be used within a decision support system which can also provide automated sedation and analgesia in

the ICU [60]. In order to achieve closed-loop sedation control in the ICU, a quantifiable feedback signal is required that reflects some measure of the patient's agitation. A non-subjective agitation assessment algorithm can be a key component in developing closed-loop control algorithms for ICU sedation. In this dissertation, a method is introduced to automatically detect pain and agitation in nonverbal patients, specifically in sedated patients in the ICU, using their facial expressions.

In clinical intensive care unit practice sedative/analgesic agents are titrated to achieve a specific level of sedation. In general, the goal of the clinician is to find the drug dose that maintains the patient at a sedation score corresponding to a moderately sedated state. This is typically done empirically, administering a drug dose that usually is in the effective range for most patients, observing the patient's response, and then adjusting the dose accordingly. However, the response of patients to any drug dose is a reflection of the pharmacokinetic and pharmacodynamic properties of the drug and the specific patient.

As part of this research, we use *pharmacokinetic* and *pharmacodynamic* modeling to find an optimal drug dosing control policy to drive the patient to a desired sedation score. Specifically, we model the pharmacokinetics and pharmacodynamics of a general sedative agent using a hybrid deterministic-stochastic model involving deterministic pharmacokinetics and stochastic pharmacodynamics. Then, using this hybrid model, we consider the sedative drug propofol and use nonnegative and compartmental modeling to model the drug pharmacokinetics (drug concentration as a function of time) and a stochastic process to represent the patient's sedation score and model drug the pharmacodynamics (drug effect as a function of concentration). Next, we use the aforementioned hybrid deterministic-stochastic model to develop an open-loop optimal control policy for ICU sedation.

Nonnegative dynamical systems involve dynamic states whose values are nonneg-

ative [48,62,89,91]. A subclass of nonnegative dynamical systems are compartmental systems [62,83,85]. Compartmental systems involve dynamical models that are characterized by conservation laws (e.g., mass, energy, fluid, etc.) capturing the exchange of material between coupled macroscopic subsystems known as compartments. These models are widespread in biological, physiological, and ecological sciences as well as engineering systems such as queuing, large-scale, telecommunications, transportation, power, and network systems, to cite but a few examples (see [48,62] and the references therein). Since nonnegative and compartmental systems have specialized structures, special control law strategies need to be developed that guarantee that the trajectories of the closed-loop plant system states remain in the nonnegative orthant of the state space for nonnegative initial conditions. In addition, for certain applications of nonnegative systems, such as active control for clinical pharmacology, we require the control (source) inputs to be nonnegative.

In this research, we use linear matrix inequalities to develop $\mathcal{H}_2$ and mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ (sub)optimal estimators and controllers for nonnegative dynamical systems. Specifically, we formulate a series of generalized eigenvalue problems subject to a set of linear matrix inequality constraints for designing $\mathcal{H}_2$ suboptimal estimators, static controllers, and dynamic controllers for nonnegative dynamical systems. The resulting $\mathcal{H}_2$ suboptimal controllers guarantee that the closed-loop plant system states remain in the nonnegative orthant of the state space.

Atrial fibrillation, a cardiac arrhythmia characterized by unsynchronized electrical activity in the atrial chambers of the heart, is a rapidly growing problem in modern societies. Electrical cardioversion and antiarrhythmic drugs are used to manage this condition, but suffer from low success rates and involve major side effects [24, 39, 47, 52]. In an alternative treatment, known as *catheter ablation*, specific parts of the left atrium are targeted for radio frequency ablation using an intracardiac catheter [86].

Application of radio frequency energy to the cardiac tissue causes thermal injury (lesions), which in turn results into scar tissue. Successful ablation can eliminate, or isolate, the problematic sources of electrical activity and effectively cure atrial fibrillation.

Magnetic resonance imaging (MRI) has been used for both pre- and and post-ablation assessment of the atrial wall [106]. MRI can aid in selecting the right candidate for the ablation procedure and assessing post-ablation scar formations. Image processing techniques can be used for automatic segmentation of the atrial wall, which facilitates an accurate statistical assessment of the region. As a first step towards the general solution to the computer-assisted segmentation of the left atrial wall, in this dissertation we propose a shape-based image segmentation framework to segment the endocardial wall of the left atrium.

1.1. Brief Outline of the Dissertation

The contents of the dissertation are as follows. In Chapter 2, we develop rule-based as well as probabilistic expert systems for cardiopulmonary management and intensive care unit sedation. In Chapter 3, we use the relevance vector machine classification technique to distinguish pain from non-pain as well as assess pain intensity levels. Then, in Chapter 4, we consider facial expression recognition using an unsupervised learning framework. Specifically, we show that different facial expressions reside on distinct subspaces if the manifold of facial images is unfolded. In particular, semi-definite embedding is used to reduce the dimensionality and unfold the manifold of facial images. Next, generalized principal component analysis is used to fit a series of subspaces to the data points and associate each data point to a subspace. In Chapter 5, we use pharmacokinetic and pharmacodynamic modeling to find an optimal drug dosing control policy to drive the patient to a desired sedation score.

Next, in Chapter 6, we formulate a series of generalized eigenvalue problems subject to a set of linear matrix inequality constraints for designing $\mathcal{H}_2$ and mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ suboptimal estimators, static controllers, and dynamic controllers for nonnegative dynamical systems. In Chapter 7, we use shape learning and shape-based image segmentation to identify the endocardial wall of the left atrium in the delayed-enhancement magnetic resonance images. Finally, in Chapter 8, we discuss ongoing research and future extensions of the research.

Chapter 2

Clinical Decision Support and Closed-Loop Control for Cardiopulmonary Management and Intensive Care Unit Sedation Using Expert Systems

2.1. Introduction

Sedation of mechanically ventilated patients in the intensive care unit is an important and challenging problem with ethical, clinical, and financial implications. At the ethical level, we have a self-evident moral imperative to provide adequate anxiolysis and analgesia for patients in the intensive care unit. From the clinical perspective, it is important that this be done without either overdosage or underdosage as either may have undesirable clinical effects. At the financial level, sedation of patients in the intensive care unit requires large investments of health care provider time, with a commensurate financial cost, while inefficient titration of sedation and analgesia may prolong intensive care unit length of stay.

While physicians select the agent(s) used for sedation, the actual administration of these agents is the responsibility of the nursing staff. The intensive care unit nurse has one of the most task-laden jobs in medicine and titration of the sedative drug dose to achieve the optimal levels of sedation can be a difficult and time consuming task. If clinical decision support systems and closed-loop control systems could be developed

for critical care monitoring and the administration of sedation, the intensive care unit nurse could be released from the intense monitoring of sedation, allowing her/him to focus on other critical tasks.

In clinical practice the dose of sedative agent is varied, or titrated, to achieve the desired level of sedation. The level of sedation is currently based on clinical scoring systems. One example is the Motor Activity Assessment Score (MAAS) [40] in which patients are given an integer score of 0–6 as follows: 0 - unresponsive; 1 - responsive only to noxious stimuli; 2 - responsive to touch or name; 3 - calm and cooperative; 4 - restless and cooperative; 5 - agitated; and 6 - dangerously agitated.

To implement closed-loop control in an acute environment, control of cardiovascular function needs to also be addressed along with sedation since hemodynamic management and control of consciousness are interrelated. For example, a major side effect of cardiac surgery is that patients can become hypertensive [99], requiring treatment to prevent cardiac dysfunction, pulmonary edema, myocardial ischemia, stroke, and bleeding from fragile sutures. Although drugs are available for treating postoperative hypertension, titration of these drugs to regulate blood pressure is often difficult. Underdosing leaves the patient hypertensive, whereas overdosing can reduce the blood pressure to levels associated with shock.

Although blood pressure control is important, cardiovascular function involves several other important variables, all of which are interrelated [99]. The intensive care unit clinician must ensure not only that blood pressure is within appropriate limits but also that *cardiac output* (i.e., the amount of blood pumped by the heart per minute) is acceptable and that the heart rate is within reasonable limits. Mean arterial blood pressure is proportional to cardiac output, with the proportionality constant denoting the systemic vascular resistance, in analogy with Ohm's law. Cardiac output is equal to the product of heart rate and *stroke volume*, the volume of blood pumped with each

beat of the heart. Stroke volume, in turn, is a function of *contractility*, the intrinsic strength of the cardiac contraction; *preload*, the volume of blood in the heart at the beginning of the contraction; and *afterload*, the impedance to ejection by the heart.

The intensive care unit clinician must balance all of these variables. Inotropic agent drugs, that is, drugs that increase the strength of contraction of the heart, also have variable effects on heart rate and afterload. There are also *vasopressor* drugs, which increase afterload, and *vasodilator* drugs, which decrease afterload. Finally, stroke volume can be improved by giving the patient intravenous fluids and increasing preload. However, too much fluid can potentially be deleterious by impairing pulmonary function as fluid builds up in the lungs. The fact that closed-loop control of blood pressure has not been widely adopted by clinicians is not surprising when one considers the complex interrelationships among hemodynamic variables.

Since cardiovascular and central nervous system functions are critical in the acute care environment, technologies have evolved for their measurements. The challenge for extending feedback control technology to the problem of sedation of critically ill patients, however, is finding the appropriate performance variable for control. Hence, the first step in the development of closed-loop control of sedation is the discovery of an objective, continuously-measurable parameter that correlates with clinician assessment of the level of sedation. Once such a parameter is discovered and validated, it then becomes necessary to use the measure of sedation for the titration of drug dose. One particularly attractive strategy is to utilize the knowledge and experience of skilled clinicians, capturing explicitly the rules expert clinicians use to decide on how to titrate drug doses depending on the level of sedation. In this chapter, we consider a rule-based expert system approach for cardiopulmonary management and intensive care unit sedation. Furthermore, we use probability theory to quantify system uncertainty and extend the proposed rule-based expert system to deal with more

realistic situations.

The contents of the chapter are as follows. First, we begin with a description of the closed-loop architecture of the sedation control system in Section 2.2. Next, in Section 2.3, we give a brief discussion on instrumentation for clinical pharmacology. In Section 2.4, we give a brief review of knowledge-based systems, and, in particular, expert systems. A rule-based expert system for cardiopulmonary management and ICU sedation control is given in Section 2.5. In Section 2.6, a rule-based expert system for respiratory management of patients in the ICU is presented. In Section 2.7, we develop a probabilistic expert system using Bayesian networks for cardiovascular function and ICU sedation control. Next, in Section 2.8 we use the expert system framework presented in Section 2.7 to design an alarm algorithm for agitation detection in the ICU. Finally, in Section 2.9, an illustrative numerical example of a probabilistic expert system for ICU sedation control is given.

2.2. Closed-Loop Sedation Control Architecture

In this section, we present a closed-loop feedback expert system architecture for ICU sedation control. A feedback system consists of an interconnection of two systems, a *forward loop* system and a *feedback loop* system. The forward loop system is driven by an input and produces an output that serves as the input to the feedback loop system. The output of the feedback loop system, in turn, serves as the input to the forward loop system. The closed-loop system is composed of the controller, the plant (patient), and the plant output measurement block (i.e., sedation assessment block). Within our sedation control framework, the plant (patient) is a dynamical system with unknown dynamics, where the input is the sedative drug dose and the output is the *patient behavior*. Patient behavior refers to patient's level of sedation and analgesia, manifested through facial expression, gross motor movement, pain, ag-

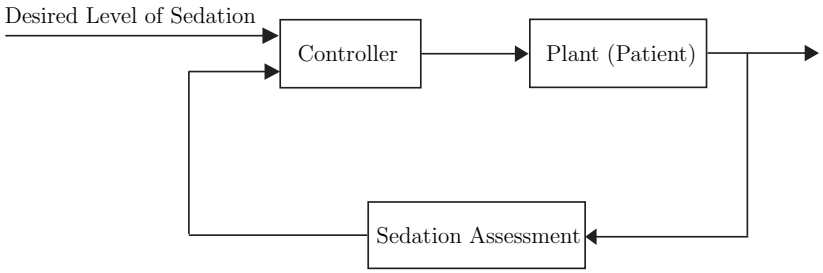

Figure 2.1: Closed-loop sedation control architecture

itation, blood pressure, and heart rate. The goal of the sedation assessment feedback block is to monitor the patient's behavior, and objectively assess the sedation level based on one of the clinical scoring systems (e.g., MAAS). The input to the controller is the desired level of sedation, and the objective assessment of sedation provided by the sedation assessment block. The closed-loop system is shown in Figure 2.1. The current clinical practice in the ICU involves human expert assessment of patient's level of sedation (corresponding to the sedation assessment block), and titration of the correct dose of sedatives (corresponding to the controller).

Closed-loop control of intensive care unit sedation is virtually nonexistent in the literature. However, control algorithms have been developed, simulated, and implemented for the related problem of closed-loop control of general anesthesia. The first of these have focused on the control of inhalation anesthesia and several adaptive control algorithms have been proposed; see [66, 71, 87, 124, 133, 147, 150] and the references therein. These algorithms have been shown to provide superior control of general inhalation anesthesia in simulations and animal studies. However, they are not directly relevant to the specific problem of ICU sedation since the controlled variable is the end-tidal anesthetic concentration. It is not possible with current technology to rapidly measure the plasma concentration of the intravenously administered drugs commonly used for ICU sedation. Thus drug concentration is not a viable control variable. Furthermore, drug concentration, even if it could be mea-

sured rapidly, is not the best control variable. We are far more interested in drug effect than concentration.

One approach to closed-loop control of sedation is to design a system that processes the information currently used by the medical staff and mimics the human process of decision making for ICU sedation. Such a system can be equipped with various sensors, including the bispectral index (a derivative of the electroencephalogram (EEG) signal) monitor [56, 128], actigraph (accelerometer for measuring hand and leg movement) [57,149], and digital imaging (for measuring facial expression and gross motor movement) [6, 32, 60]. In a recent study, machine learning methods have been used to assess the level of pain in patients using facial expressions and analyze the correlation between computer and human expert pain intensity assessments [53, 54]. With measurements provided by different viable sensors, an expert system can be designed which mimics expert human actions and follows a similar decision making process.

2.3. Instrumentation for Clinical Pharmacology

The sensors used in the intensive care unit to monitor patient status include those that measure hemodynamic status, respiratory status, renal function, and central nervous function. Hemodynamic status is most typically assessed by continuous monitoring of heart rate and electrocardiograph (ECG). The ECG measures the electrical potential difference between skin electrodes placed at various sites on the torso and limbs, and can be analyzed to provide continuous heart rate measurement as well as identify signs of cardiac dysfunction. Hemodynamic function is also assessed using blood pressure measurements. While this may be done using noninvasive methods, it is most typically done by placing a small plastic catheter directly into an artery (most often the radial artery as it passed through the underside of the wrist) and

then using a pressure transducer to convert the pulse pressure wave into an electrical signal. In a similar fashion, catheters are also often placed into large central veins (such as the internal jugular vein) so that their tips are situated close to the entry of the main veins (superior vena cava or inferior vena cava) returning blood to the heart. Pressure waves in these veins are then transduced into electrical signals to provide the central venous pressure. This gives an indirect measure of the volume of blood in the heart which is a major determinant of cardiac output, the volume of blood pumped by the heart per minute.

In some situations in which there is more profound cardiac dysfunction, a pulmonary artery catheter is placed. This is a catheter that runs through the heart into the pulmonary artery (i.e., the artery going from the heart to the lungs) and can measure pressures in the pulmonary artery (another indirect measure of volume in the heart) as well as directly measure cardiac output. Finally, it is important to monitor the adequacy of blood flow to the various tissues of the body. One common technique is to measure the amount of oxygen in venous blood. If the delivery of oxygen to tissue decreases, then there will be a greater relative extraction of oxygen from the delivered blood by the tissue, and hence, the venous blood returning to the heart will have less oxygen in it. This is most typically measured as the percentage of hemoglobin molecules (the primary carrier of oxygen in the blood) that are bound to oxygen (referred to as the venous saturation).

The purpose of respiration is to eliminate carbon dioxide from and deliver oxygen to the blood. Hence, the most important monitors of respiratory function are measures of carbon dioxide and oxygen in the blood. With the most commonly used sensor technologies these are not directly measurable; however, it is possible to continuously measure hemoglobin oxygen saturation, the percentage of hemoglobin in arterial blood that is bound to oxygen, using absorbance spectroscopy and light emit-

ting diode technology. In addition, many intensive care units use continuous analysis of gas exhaled from the lungs to measure end-tidal carbon dioxide concentration, an indirect and approximate measure of blood carbon dioxide concentrations. Furthermore, modern mechanical ventilators are equipped to measure the pressure used to expand the lungs when the patient is undergoing mechanical ventilation, as well as respiratory rate.

Assessment of renal function is not as sophisticated as either hemodynamic or respiratory monitoring. Currently renal function is most typically assessed by the continuous measurement of urine output. Sensors for assessment of central nervous system function are currently in their infancy, at least as far as routine clinical use is concerned.

2.4. A Brief Review of Expert Systems

In this section, we review knowledge-based systems, and, in particular, expert systems. A knowledge-based system is a computer program that is capable of making deductions based on the information provided by the user and the information stored in its knowledge base. In other words, a knowledge-based system is a system which applies a "rules of thumb" approach to a symbolic representation of knowledge [80]. The main characteristic that distinguishes a knowledge-based system from a conventional computer program is its structure [74]. In conventional computer programs, the knowledge and the computational/analytical components of the program are coupled. Knowledge-based systems, however, have two main independent components; namely, the *knowledge base*, which stores the information, and the *inference engine*, which makes assertions based on the available knowledge. Expert systems are a subclass of knowledge-based systems, where their objective is to emulate the human expert behavior [29, 74].

Expert systems in general deal with two different types of problems: deterministic versus stochastic. As a result, expert systems belong to one of the two general classes of *i*) deterministic expert systems and *ii*) stochastic expert systems. Deterministic expert systems are also referred to as *rule-based expert systems* due to the fact that in such systems the deduction process is based on a series of rules [29]. A more challenging set of problems is that involving uncertainty in knowledge and in the problem variables. Stochastic expert systems specifically deal with such problems and different frameworks exist to address uncertainty including certainty factors [26], fuzzy logic [154], theory of evidence [130], and, more recently, probability theory [29]. In the probabilistic approach, a joint probability distribution function over the set of variables is defined and the inference is based on probability rules. Such expert systems are referred to as *probabilistic expert systems*.

The most basic element of an expert system is knowledge. Knowledge, in general, can be categorized into *abstract* and *concrete*. Abstract knowledge refers to the knowledge stored in the permanent memory of the system (i.e., the knowledge base), which includes rules and probability distributions. Concrete knowledge refers to information particular to a specific case and situation, which is recorded in the working memory of the system (e.g., a set of observed symptoms in a medical diagnosis problem, which is different from patient to patient).

The components of a general expert system are outlined as follows.

i) *Human component*. In the design of an expert system, the *subject-matter expert* (or simply, the *expert*) collaborates with a *knowledge engineer*, where the expert provides the knowledge and the knowledge engineer transfers the knowledge into a structured form which allows further processing by a computer.

ii) *Knowledge base*. The knowledge base consists of structured knowledge (i.e.,

knowledge provided by the expert which is translated by the knowledge engineer to a computer-friendly language), and the relationships among such structured knowledge.

iii) *Knowledge and information acquisition subsystem.* This component is responsible for the in-flow of information from the expert into the system, and solicits knowledge if the initial knowledge is limited and no conclusion can be made. It also has the capability to add new knowledge to the knowledge base. The human user is usually the source of additional knowledge.

iv) *Coherence control.* This component, which has been included in expert systems only recently, is responsible for maintaining consistency in the knowledge base in order to avoid incoherence. Even human experts could provide contradictory statements, especially when dealing with complex problems. To avoid absurd conclusions (e.g., the probability of an event being greater than 1) resulting from inconsistencies in the knowledge base, the coherence control component checks the existing knowledge and the knowlege being added to the knowledge base for any inconsistencies and informs the human user if a contradiction is found. It is worth noting that a similar idea exists for identifying discrepancies in a software implementation of a general control system. More specifically, dynamical systems theory is used to identify the allowable domain of operation for each state, and any violation of these constraints is regarded as an inconsistency. For further details, see [49].

v) *Inference engine.* This is the main component of the expert system and serves to apply abstract knowledge to concrete knowledge and draws conclusions.

vi) *User interface.* This component serves as the medium between the human user and the expert system.

vii) *Action execution subsystem.* This component is responsible for performing actions and executing commands specifically in expert systems which are designed to take control actions.

viii) *Explanation subsystem.* This component explains the actions taken by the action execution subsystem or conclusions reached by the inference engine. This component clarifies the logical deduction process for the human user.

ix) *Learning subsystem.* This component is responsible for acquiring new knowledge (learning new rules, finding a better estimate of system parameters, etc.) based on new data and exposure to new situations of the problem. As a result, the performance of an expert system possessing this component is expected to improve as the number of new situations to the system increases.

An expert system might possess only a subset of the aforementioned components. For a more comprehensive discussion on expert systems, see [29].

Next, we introduce the rule-based expert system, which deals with deterministic problems. In order to solve deterministic problems, a set of *rules* can be used to define the relationships among a set of objects. A rule is defined as a logical statement defining a relationship between the *premise* and the *conclusion*. The *abstract knowledge* in rule-based expert systems is given by rules and concrete knowledge is given by *facts*. A fact can be regarded as a special rule in which the premise is always true. Different inference strategies are used in rule-based expert systems. In this chapter, we use *modus ponens*, the most commonly used strategy, which is a basic element of classical logic [29]. In classical logic, each statement is either *true* or *false*. Based on the *modus ponens* strategy, given a set of facts, the premise of all the rules is examined and the true value is assigned to the conclusion of a rule for which its premise has the true value. A complementary inference strategy is the *modus tollens*,

where the negation of the conclusions are examined, and if true, then the negation of its premise is assigned the true value.

2.5. A Rule-Based Expert System for Cardiopulmonary Management and ICU Sedation Control

In this section, we introduce a simple rule-based expert system for cardiopulmonary management and ICU sedation control. We assume a sedation protocol with the drugs *propofol* (as the primary agent) and *fentanyl* (as the secondary agent) with sedation assessment using the MAAS scale to illustrate a rule-based system for control of ICU sedation. Before stating the rules, we give a short description of the drugs used in the ICU sedation and their effects. For complete details on the use of sedatives and analgesics in the ICU, see [81]. In addition, we show how this expert system can activate a secondary expert system to regulate patient hemodynamics.

Propofol, or *2,6-diisopropylphenol*, is an intravenous hypnotic agent that in low doses can produce anxiolysis and in higher doses, hypnosis (i.e., lack of responsiveness and lack of consciousness). Propofol is widely used for ICU sedation because of this spectrum of pharmacodynamic effects and also because of its pharmacokinetics. It is typically administered as a continuous infusion and it is a short acting drug that can be readily titrated, that is, if the infusion rate is increased the blood level increases relatively quickly. Hence, the pharmacological effect of the drug can be quickly varied by varying the infusion rate.

While propofol has primary pharmacodynamic and pharmacokinetic effects that suit it well for ICU sedation, there are some serious side effects that may limit its usefulness. It causes dilation of both arteries and veins, as well as mild depressant effects on the heart, that can cause in turn excessive drops in blood pressure. Furthermore, it does not have analgesic effects, and thus, is ineffective in treating pain.

Since pain often results in increases in heart rate and blood pressure, propofol can be paradoxically associated with either hypotension (at excessive doses) or hypertension (when the patient has untreated pain).

Because of the hypotensive effects of propofol, and since it does not treat pain-induced hypertension, our rule-based system also uses fentanyl as a secondary agent for sedation. Fentanyl is a synthetic opioid and potent analgesic. It can be quite effective in the treatment of pain-induced hypertension. At the same time, it has mild sedative effects (although even in high doses it does not reliably produce hypnosis). Since it does not have the pronounced hypotensive effects of propofol, it can be used for its sedative effects in hypotensive patients. While it can be employed as a primary agent for sedation, we will illustrate our rule-based system by assuming it is a second-line agent. This is motivated largely by the fact that it is not quite as fast-acting as propofol and not as easily titrated.

The rule base of a simple ICU sedation control expert system is summarized in Table 2.1. The desired level of sedation corresponds to an MAAS score of 3. The premise of each rule involves the current (M) and previous (M') MAAS scores, blood pressure (BP), and heart rate (HR). The conclusion of each rule consists of primary action and secondary action. The required dose of drugs, denoted by primary action in the table, is given in the first part of the conclusion of each rule. The symbols "$\uparrow$" and "$\downarrow$" denote increase and decrease in the infusion rate of the drug, respectively. Furthermore, "+/ $\uparrow$ fentanyl" stands for "if the patient is already on fentanyl, then increase the fentanyl infusion rate by 1 mcg/kg/hr after a 2 mcg/kg bolus dose and, if not, then start fentanyl at 1 mcg/kg/hr after a 2 mcg/kn bolus dose." Finally, note that we have assumed that for a given MAAS score, the previous MAAS score is within its ± 1 range. This is not a limiting assumption if the sedation assessment is performed frequently so that we capture the dynamics of the MAAS score.

Table 2.1: The rule base of a simple ICU Sedation control expert system, which involves the current MAAS score (M), previous MAAS score (M'), and patient's blood pressure (BP) and heart rate (HR).

M	M'	BP/HR	Primary Action	Secondary Action
0 & 1	-	BP$\geq$ 150 or HR$\geq$ 120	discontinue fentanyl & propofol	activate HDCES
		90 <BP< 150	discontinue fentanyl & propofol	-
		BP$\leq$ 90	discontinue fentanyl & propofol	activate HDCES
2	1	BP$\geq$ 150 or HR$\geq$ 120	+/$\uparrow$ fentanyl	if on fentanyl activate HDCES
		90 <BP< 150	-	-
		BP$\leq$ 90	25% $\downarrow$ propofol	-
	2	BP$\geq$ 150 or HR$\geq$ 120	25% $\downarrow$ propofol, +/$\uparrow$ fentanyl	if on fentanyl activate HDCES
		90 <BP< 150	25% $\downarrow$ propofol	-
		BP$\leq$ 90	25% $\downarrow$ propofol	-
	3	BP$\geq$ 150 or HR$\geq$ 120	25% $\downarrow$ propofol, +/$\uparrow$ fentanyl	if on fentanyl activate HDCES
		90 <BP< 150	25% $\downarrow$ propofol	-
		BP$\leq$ 90	50% $\downarrow$ propofol	-
3	2	BP$\geq$ 150 or HR$\geq$ 120	+/ $\uparrow$ fentanyl	-
		90 <BP< 150	-	-
		BP$\leq$ 90	-	activate HDCES
	3	BP$\geq$ 150 or HR$\geq$ 120	+/ $\uparrow$ fentanyl	-
		90 <BP< 150	-	-
		BP$\leq$ 90	-	activate HDCES
	4	BP$\geq$ 150 or HR$\geq$ 120	+/ $\uparrow$ fentanyl	-
		90 <BP< 150	-	-
		BP$\leq$ 90	-	activate HDCES
4	3	BP$\geq$ 150 or HR$\geq$ 120	50% $\uparrow$ propofol	-
		90 <BP< 150	25% $\uparrow$ propofol	-
		BP$\leq$ 90	25% $\downarrow$ propofol, +/ $\uparrow$ fentanyl	-
	4	BP$\geq$ 150 or HR$\geq$ 120	50% $\uparrow$ propofol	-
		90 <BP< 150	25% $\uparrow$ propofol	-
		BP$\leq$ 90	25% $\downarrow$ propofol, +/ $\uparrow$ fentanyl	-
	5	BP$\geq$ 150 or HR$\geq$ 120	+/ $\uparrow$ fentanyl	-
		90 <BP< 150	-	-
		BP$\leq$ 90	-	activate HDCES
5	4	BP$\geq$ 150 or HR$\geq$ 120	50% $\uparrow$ propofol, +/ $\uparrow$ fentanyl	-
		90 <BP< 150	50% $\uparrow$ propofol	-
		BP$\leq$ 90	+/ $\uparrow$ fentanyl	activate HDCES
	5	BP$\geq$ 150 or HR$\geq$ 120	50% $\uparrow$ propofol, +/ $\uparrow$ fentanyl	-
		90 <BP< 150	50% $\uparrow$ propofol	-
		BP$\leq$ 90	+/ $\uparrow$ fentanyl	activate HDCES
	6	BP$\geq$ 150 or HR$\geq$ 120	25% $\uparrow$ propofol, +/ $\uparrow$ fentanyl	-
		90 <BP< 150	25% $\uparrow$ propofol	-
		BP$\leq$ 90	+/ $\uparrow$ fentanyl	activate HDCES
6	-	BP$\geq$ 150 or HR$\geq$ 120	100% $\uparrow$ propofol, +/ $\uparrow$ fentanyl	-
		90 <BP< 150	100% $\uparrow$ propofol	-
		BP$\leq$ 90	+/ $\uparrow$ fentanyl	activate HDCES

Table 2.2: The rule base of a simple hemodynamics control expert system, which involves the patient's blood pressure and heart rate.

BP	HR	Action
BP$\geq$ 150	HR$\geq$ 110	Beta-Blocker
	HR$<$ 110	Vasodilator
90 $<$BP$<$ 150	HR$\geq$ 110	Beta-Blocker
	HR$<$ 110	-
BP$<$ 90	-	See Hypotension Protocol

The second part of the conclusion of each rule, denoted by secondary action, involves activating a secondary expert system, namely, the hemodynamic control expert system (HDCES). Depending on the blood pressure and the heart rate of the patient, the ICU sedation control expert system can activate the hemodynamic control expert system to regulate patient cardiovascular function. The rule base of a hemodynamic control expert system is summarized in Table 2.2.

The most obvious monitor of cardiovascular function in the intensive care unit is blood pressure, and treatment of blood pressure is a very common activity in the intensive care unit. The hemodynamic expert system we present is based on the treatment of blood pressure. To a first approximation, the circulation can be described as a very simple direct-current system conforming with a hemodynamic version of Ohm's law. Specifically, we can describe the relationship between mean arterial blood pressure (the hemodynamic equivalent of voltage) and cardiac output (the hemodynamic equivalent of current) by [2]

$$\mathrm{MAP}(t) = \mathrm{CO}(t) \times \mathrm{SVR}(t) + \mathrm{CVP}(t), \quad t \geq 0, \tag{2.1}$$

where $\mathrm{MAP}(t)$ is mean arterial blood pressure, $\mathrm{CO}(t)$ is cardiac output (the volume of blood the heart pumps per minute), $\mathrm{SVR}(t)$ systemic vascular resistance (an index of arteriolar compliance or constriction throughout the body), and $\mathrm{CVP}(t)$ is central venous pressure (the venous pressure of the right atrium of the heart). Since $\mathrm{CVP}(t)$

is typically much less significant than $\mathrm{CO}(t) \times \mathrm{SVR}(t)$, we can see that any analysis of blood pressure perturbation should focus on whether the change in blood pressure is due to a change in cardiac output or a change in systemic vascular resistance.

For *hypotensive* patients (systolic blood pressure $< 90\,\mathrm{mm\ Hg}$ or mean blood pressure $< 60\,\mathrm{mm\ Hg}$) the expert system algorithm first entails an evaluation of cardiac output. This can be done by direct measurement or by indirect means. There are a number of technologies for the measurement of cardiac output, including thermodilution, lithium dilution, and analysis of the contour of the arterial pulse wave. Since these technologies are not that frequently employed, the adequacy of cardiac output is often assessed by measurement of central venous hemoglobin oxygen saturation as described in Section 2.3. In some situations it is necessary to assess the adequacy of cardiac output using clinical findings such as poor peripheral circulation, acidosis, or poor urine output.

Returning to the basic equation for hemodynamics (2.1), if cardiac output is inadequate, then efforts to correct hypotension should be directed toward improving cardiac output. Cardiac output equals stroke volume (i.e., the amount of blood pumped by the heart each time it beats) multiplied by heart rate. If the heart rate is exceptionally low, then one can administer drugs to speed up the heart. More frequently, the focus is on increasing stroke volume. Stroke volume is determined by *preload*, the term used in the hemodynamic literature to refer to the amount of blood volume in the heart at the onset of each contraction, *contractility*, the strength of the contraction, and, to a lesser degree, *afterload*, roughly the load the heart faces in order to pump blood. The first step is to ensure adequate preload. This is evaluated by consideration of the central venous pressure, or pulmonary artery wedge pressure if a pulmonary artery catheter is in place, or by analysis of how the peak systolic arterial pressure changes with inspiration if the patient is undergoing mechanical ventilation,

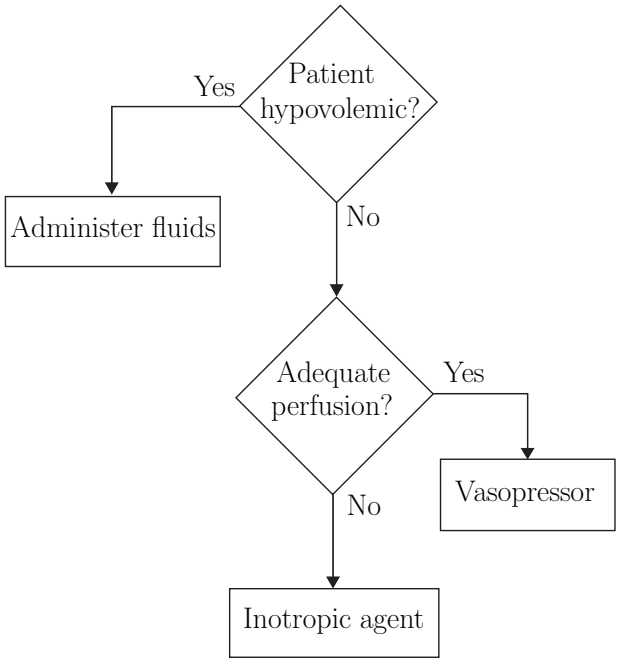

Figure 2.2: Hypotension protocol flow chart.

by using echocardiography to visualize the heart, or by simply giving the patient a bolus of intravenous fluids and observing the blood pressure response. If preload is adequate but stroke volume is assessed as inadequate, then the only recourse is to administer drugs (positive inotropes) that increase the contractility of the heart.

In many situations, especially in patients with infections, the cardiac output is adequate or higher than normal but the blood pressure is still low. Referring again to the basic equation of hemodynamics (2.1), one must conclude that systemic vascular resistance is low. In this case, we administer drugs (vasopressors) that increase systemic vascular resistance. The hypotension protocol flow chart is given in Figure 2.2.

The treatment of *hyerptension* (systolic blood pressure $> 150\,\text{mm Hg}$) follow somewhat similar considerations. If the patient has an elevated heart rate as well as an elevated blood pressure, the usual cause is increased contractility. Since high heart rates potentially can cause *myocardial ischemia* (i.e., inadequate matching of oxygen

delivery to the heart to the demand) it is most appropriate to treat the patient with beta-blockers, that is, drugs that decrease both heart rate and contractility. However, if the heart rate is not elevated, then the most likely cause of the hypertension is elevated systemic vascular resistance and the best treatment would be vasodilaotors, that is, drugs that decrease systemic vascular resistance.

2.6. A Rule-Based Expert System for ICU Respiratory Management

In this section, we introduce a simple rule-based expert system for ICU respiratory management. Note that the respiratory management expert system and the cardiopulmonary management and ICU sedation control expert system are two independent control systems running concurrently. In respiratory management, the goal is to control the arterial partial pressure of CO_2 (carbon dioxide) denoted by $P_{\rm a}CO_2(t)$ and the pH of arterial blood. The means to do this are embodied in two equations; one relating $P_{\rm a}CO_2(t)$ to *alveolar ventilation* (i.e., the volume of gas exchange in the lungs in a given unit of time), and the other, the Henderson-Hasselbalch equation [104], relating blood pH to $P_{\rm a}CO_2(t)$ and the concentration of bicarbonate in the blood denoted by $[HCO_3^-](t)$.

The relationship between $P_{\rm a}CO_2(t)$ and ventilation is given by [104]

$$P_{\rm a}CO_2(t) = 0.863\frac{VCO_2}{V_{\rm a}(t)}, \quad t \geq 0, \tag{2.2}$$

where VCO_2 is the total body production of CO_2 per minute and is approximately 259 ml/min in healthy subjects, 0.863 is a constant to reconcile units, and $V_{\rm a}(t)$ is alveolar ventilation. In patients who are totally dependent on mechanical ventilation (and not taking any independent breaths) $V_{\rm a}(t)$ is given by [104]

$$V_{\rm a}(t) = (TV(t) - V_{\rm d})RR(t), \quad t \geq 0, \tag{2.3}$$

where $TV(t)$ denotes the volume of each breath set on the ventilator, $RR(t)$ denotes the respiratory rate set on the ventilator, and V_{d} denotes the dead space of the lungs. The product $TV(t)RR(t)$ is referred to as the *minute ventilation* [104] and V_{d} is approximately 1/3 of minute ventilation in healthy subjects. Note that changes in V_{a} and V_{d} are very gradual and these variables can be regarded as constants.

In actual practice, the clinical staff set the value of $TV(t) \equiv TV$ and $RR(t) \equiv RR$ on the ventilator and, using (2.3), we can see that these are the only variables that the clinician can manipulate to control $P_{\mathrm{a}}\mathrm{CO}_2(t)$. In modern critical care practice, TV is set to 6 ml/kg. As a result, the primary variable to control $P_{\mathrm{a}}\mathrm{CO}_2(t)$ is $RR(t)$. The other variables in the above equation (i.e., V_{d} and $V\mathrm{CO}_2$) are specific to the patient and her/his physiology, or rather, pathophysiology.

In addition to controlling $P_{\mathrm{a}}\mathrm{CO}_2(t)$, the clinician can control blood pH levels. In particular [104],

$$pH(t) = 6.1 + \log_e \left(\frac{[\mathrm{HCO}_3^-](t)}{0.03 P_{\mathrm{a}}\mathrm{CO}_2(t)} \right), \quad t \geq 0, \tag{2.4}$$

where $[\mathrm{HCO}_3^-](t)$ is the bicarbonate ion concentration in arterial blood and 0.03 is a constant to reconcile units. This equation reflects the fact that dissolved carbon dioxide reacts with water to form carbonic acid, which will lower the pH of blood. Once the desired $P_{\mathrm{a}}\mathrm{CO}_2(t)$ is attained by manipulating $RR(t)$ (or, less commonly, $TV(t)$), control of the desired pH can only then be achieved by manipulating $[\mathrm{HCO}_3^-](t)$. This is done by either administering bicarbonate in the case of acidosis or, less commonly, by administering an acidifying agent such as the drug *acetazolamide* or dilute hydrochloric acid. Since the deleterious effects of acidosis are more immediate and readily apparent than alkalosis, most clinicians administer acidifying agents only within specific clinical contexts.

The first step in respiratory management involves measuring the arterial gas.

This is done intermittently by taking a small sample of blood from an artery and sending it to a laboratory where the partial pressure of carbon dioxide in the blood $P_{\mathrm{a}}CO_2(t)$ is measured using electrochemical methods. In many clinical settings, the arterial $P_{\mathrm{a}}CO_2(t)$ can be approximated by *end-tidal* CO_2, the concentration of CO_2 in exhaled gas at the end of expiration. This is conveniently measured using near infrared spectroscopy of exhaled gas collected by sampling from the endotracheal tube, the interface between the patient and the mechanical ventilator. This can be done on a breath-by-breath basis.

In most cases involving ventilation control the primary goal is to normalize pH so that $P_{\mathrm{a}}CO_2(t)$ is controlled to facilitate achieving a normal pH. However, in the case of increased *intracranial pathology* it is important to maintain a normal carbon dioxide level as well as normal pH level. The brain is enclosed in a closed vault (i.e., the skull). If the brain becomes *edematous* (i.e., excessive accumulation of serous fluid), then this will increase the pressure (the intracranial pressure) inside this closed vault. If the intracranial pressure becomes too great, then the brain will be compressed and this can result in serious injury if not death. In cases of intracranial pathology (e.g., brain tumors, traumatic injury to the brain, and bleeding in the brain) there will be increased edema, and hence, increased intracranial pressure. This is exacerbated by increased carbon dioxide as this increases blood flow to the brain and increases the edema fluid load.

On the other hand, a markedly decreased carbon dioxide can lower cerebral blood flow and, if severe, can result in *cerebral ischemia* (i.e., inadequate blood flow to the brain). Hence, it is important to not only control $pH(t)$ but also $P_{\mathrm{a}}CO_2(t)$ in patients with intracranial pathology who require mechanical ventilation. The rule base of a simple respiratory management expert system is summarized in Table 2.3.

Table 2.3: Rule base of a simple respiratory management expert system. (a) Absence of intracranial pathology. (b) Presence of intracranial pathology.

(a)

pH	Action
$pH < 7.32$	$\uparrow RR$, wait 30 min., repeat arterial blood gas measurement
$7.32 < pH < 7.45$	-
$pH > 7.45$	$\downarrow RR$, wait 30 min., repeat arterial blood gas measurement

(b)

pH	$P_{\mathrm{a}}CO_2$	Action
$pH < 7.32$	$P_{\mathrm{a}}CO_2 > 40$	Increase RR, wait 30 min., repeat arterial blood gas measurement
	$30 \leq P_{\mathrm{a}}CO_2 \leq 40$	Administer $[HCO_3^-]$ per base deficit
	$P_{\mathrm{a}}CO_2 < 30$	$\downarrow RR$, administer $[HCO_3^-]$
$7.32 \leq pH \leq 7.49$	$P_{\mathrm{a}}CO_2 < 30$	$\downarrow RR$, wait 30 min., repeat arterial blood gas measurement
	$30 < P_{\mathrm{a}}CO_2 < 40$	No action
	$P_{\mathrm{a}}CO_2 > 40$	$\uparrow RR$, wait 30 min., repeat arterial blood gas measurement
$pH > 7.49$	$30 < P_{\mathrm{a}}CO_2$	$\downarrow RR$, repeat arterial blood gas measurement
	$30 < P_{\mathrm{a}}CO_2 < 40$	consider acidifying agent, if given, repeat arterial blood gas measurement
	$P_{\mathrm{a}}CO_2 > 40$	$\uparrow RR$, consider acidifying agent

2.7. A Probabilistic Expert System for Cardiopulmonary Management and ICU Sedation Control

One of the limitations of the rule-based expert systems proposed in Sections 2.5 and 2.6 is their inability to deal with uncertainty. More specifically, the rule-based expert system in Section 2.5 assumes perfect accuracy in the measurement of present and previous MAAS scores, blood pressure, and heart rate. While current technology allows for high accuracy measurements of blood pressure and heart rate, the MAAS score, which quantifies the level of sedation and agitation of the patient, is subjective and can result in inconsistencies and variability in sedation administration. Moreover, in a rule-based expert system there is no uncertainty associated with the rules. A more general approach would allow for rules with multiple conclusions, where a different level of uncertainty is associated with each conclusion.

In this section, we use probability theory to quantify uncertainty to extend the rule-based expert system given in Section 2.5 to deal with more realistic situations. The rule-based respiratory management expert system given in Section 2.6 mainly

uses $P_{\mathrm{a}}\mathrm{CO}_2$ and blood pH data for decision making. The same framework can be used to construct a probabilistic expert system for respiratory management.

In the Bayesian interpretation of probability, as opposed to the classical interpretation, the probability of an event is an indication of the *uncertainty* associated with the event rather than its *frequency* [15]. In the probabilistic approach to expert systems, the system variables are regarded as random variables and, in contrast to rule-based expert systems, probabilistic expert systems do not possess "if-then" rules but rather the relationship between the variables is defined using a joint probability distribution [29]. If the joint probability distribution of a probabilistic model is known, probabilities associated with different situations can be computed using marginalization and probability conditioning [108].

A drawback of the probabilistic approach to expert systems is computational complexity. The computational complexity increases with the increase in the number of random variables and the number of possible values they can take. This increase is exponential in the number of random variables. *Bayesian networks* [50] (also known as *belief networks*) is a graphical framework in machine learning which exploits the conditional independence between variables to reduce the computational complexity of the probabilistic model.

Before stating the main results of this section, we need the following definitions.

Definition 2.1 [65, 138]. A *directed graph* $\mathfrak{G}$ is a pair $(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ is the set of *vertices* and $\mathcal{E} = \{e_1, e_2, \ldots, e_M\} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of *edges*. Every edge $e_l \in \mathcal{E}$, $l \in \{1, \ldots, M\}$, corresponds to an ordered pair of vertices $(v_i, v_j) \in \mathcal{V} \times \mathcal{V}$, where v_i and v_j are the initial and terminal vertices of the edge e_l. In this case, e_l is *incident into* v_j and *incident out* of v_i; v_i is the *parent* of v_j and v_j is the *child* of v_i. Moreover, $\Pi_w \triangleq \{v \in \mathcal{V} : (v, w) \in \mathcal{E}\}$ is the *set of all parents* of $w \in \mathcal{V}$.

A *directed path* from v_{i_1} to v_{i_k} is a set of distinct vertices $\{v_{i_1}, v_{i_2}, \ldots, v_{i_k}\}$ such that $(v_{i_j}, v_{i_{j+1}}) \in \mathcal{E}$, $j = 1, \ldots, k-1$. A directed path is *closed* if $v_{i_1} = v_{i_k}$. A *cycle* is a nontrivial closed path where all the vertices (except for the first and last) are distinct. A *directed acyclic graph* is a directed graph containing no cycles.

For the next definition, $p(\cdot)$ and $p(\cdot|\cdot)$ denote the probability density function and the conditional probability density function operators, respectively.

Definition 2.2 [50]. Let $\mathcal{V} = \{X_1, X_2, \ldots X_n\}$, where X_i, $i = 1, \ldots, n$, is a random variable and takes on values from a set $\mathcal{X}_i \subset \mathbb{R}$, $i = 1, \ldots, n$. A *Bayesian network* $\mathcal{B}$ is an ordered pair $(\mathfrak{G}, \Theta)$, where $\mathfrak{G} = (\mathcal{V}, \mathcal{E})$ is a directed acyclic graph, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, and Θ is the set characterizing the probabilistic relationship between the vertices (random variables) and is defined by

$$\Theta \triangleq \{p\left(x_i | X_{i_1} = x_{i_1}, \ldots, X_{i_M} = x_{i_M}\right) : x_i \in \mathcal{X}_i,\, i = 1, \ldots, n,\, i_1, \ldots, i_M \in \mathcal{I}_{X_i}\} \quad (2.5)$$

where $\mathcal{I}_{X_i} \triangleq \{j : X_j \in \Pi_{X_i}\}$ is the parent index set of X_i, $M = \text{ card}(\Pi_{X_i})$, and $\text{card}(\cdot)$ is the cardinality operator.

Note that a Bayesian network $\mathcal{B} = (\mathfrak{G}, \Theta)$ defines a unique joint probability distribution over $\mathcal{V}$ given by

$$p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} p\left(x_i | x_{i_1}, \ldots, x_{i_M}\right), \quad (2.6)$$

where $i_1, \ldots, i_M \in \mathcal{I}_{X_i}$, and where, for simplicity of exposition, we denote the conditional probability density function $p\left(x_i | X_{i_1} = x_{i_1}, \ldots, X_{i_M} = x_{i_M}\right)$ by $p\left(x_i | x_{i_1}, \ldots, x_{i_M}\right)$. Moreover, a Bayesian network represents the causal relationships between different random variables. More specifically, if $X_j \in \Pi_{X_i}$, $i, j \in \{1, \ldots, n\}$, then X_j directly influences (causes) X_i. This interpretation of an edge between two vertices

is crucial in the construction of the Bayesian network. Each random variable is either *observed* (i.e., its value is known), or *hidden* (i.e., its value is unknown). In the graphical representation of a Bayesian network, the vertices corresponding to the observed random variables are shaded. Given a Bayesian network and the set of observed random variables, the *inference* involves finding the posterior probability distribution of any set of random variables given the observed random variables by marginalizing the joint probability distribution. An advantage of Bayesian networks is that they reduce the computational complexity of the inference stage. For a more comprehensive discussion on Bayesian networks, see [15, 50, 108].

In this section, we use a Bayesian network framework to design a probabilistic expert system for cardiopulmonary management and ICU sedation control. We first start by constructing a Bayesian network for ICU sedation control. Next, we extend the Bayesian network to control patient hemodynamics. Let $\mathcal{B} = (\mathfrak{G}, \Theta)$ represent the Bayesian network and let the patient's current MAAS score, previous MAAS score, blood pressure, heart rate, and required drug dose for sedation be given by the random variables M, M', B, H, and D, respectively, where $\text{range}(M) = \text{range}(M') = \{0, 1, \ldots, 6\}$, $\text{range}(D) = \{1, 2, \ldots, 12\}$, $\text{range}(B) = \text{range}(H) = \mathbb{R}_+$, where $\mathbb{R}_+$ denotes the set of positive scalars, and, for a given function $f : \mathcal{X} \to \mathcal{Y}$, $\text{range}(f) \triangleq \mathcal{Y}$. Note that there are 12 distinct actions (primary action) given in the first part of the conclusion of each rule in Table 2.1, and hence, we have assigned a unique number to each distinct action. The graph $\mathfrak{G}$ for this Bayesian network is given in Figure 2.3. The current and previous MAAS scores, blood pressure, and heart rate, which constitute the inputs to the expert system and directly influence the required drug dose, are observed and their corresponding vertices are shaded in Figure 2.3.

A potential problem associated with the Bayesian network given in Figure 2.3 is its inability to capture the uncertainty associated with the measurement of the MAAS

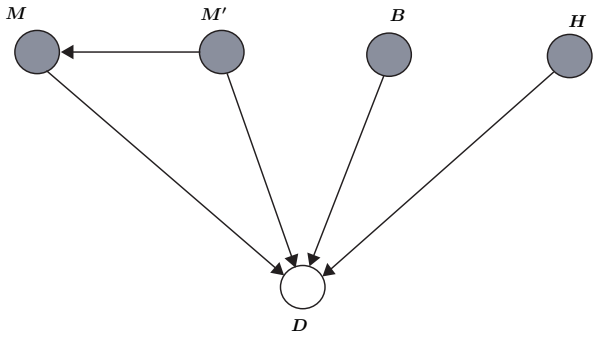

Figure 2.3: The graph of a Bayesian network capturing the relationships between the current MAAS score (M), previous MAAS score (M'), blood pressure (B), heart rate (H), and required drug dose (D).

scores. In particular, in order to perform a meaningful inference, the exact values of the current and previous MAAS scores should be known (observed). However, as discussed earlier, the assessment process is highly subjective, and the assessed scores can involve a high degree of uncertainty. A closer examination of the current and previous MAAS scores reveals that these random variables are essentially hidden variables, that is, they are "driven" by other factors. The MAAS score reflects the patient's agitation and sedation level, which can be observed through facial expressions, gross motor movement, guarding (i.e., a response in which the patient withdraws from a potentially noxious stimulus), heart rate and blood pressure stability, noncardiac sympathetic stability, and nonverbal pain scale. These observed factors can be regarded as random variables taking on values from appropriate sets. For example, machine learning techniques can be used to classify photographs based on the patient's facial expressions into pain and non-pain classes, which in turn can be used to assess pain intensity on a scale from 0 to 100 [53, 54]. Hence, a more complete model for the probabilistic expert system should include these observed random variables as well.

Let the random variables $U_1, \ldots, U_6$ represent the current objective assessment of the facial expression, gross motor movement, guarding, heart rate and blood pressure stability, noncardiac sympathetic stability, and nonverbal pain scale, respectively,

and let $U_1', \ldots, U_6'$ represent the previous objective assessment of these variables. Moreover, let range(U_1) = range$(U_1') = \{0, 1, 2\}$, where 0, 1, and 2 denote, respectively, a relaxed face, grimacing and moaning face, and grimacing and crying face; range(U_2) = range$(U_2') = \{0, 1, 2\}$, where 0, 1, and 2 denote, respectively, lying quietly, cautious movement, and restless withdrawal; range(U_3) = range$(U_3') = \{0, 1, 2\}$, where 0, 1, and 2 denote, respectively, lying quietly, splinting and tense, and rigid and stiff; range(U_4) = range$(U_4') = \{0, 1, 2\}$, where 0, 1, and 2 denote, respectively, stable, moderate change, and marked change; range(U_5) = range$(U_5') = \{0, 1, 2\}$, where 0, 1, and 2 denote, respectively, warm and dry skin, flushed and sweaty, and pale and sweaty; and range(U_6) = range$(U_6') = \{0, \ldots, 10\}$, where 0 and 10 denote, respectively, no pain and extreme pain.

The graph of the Bayesian network which includes these new random variables is given in Figure 2.4. Note that the current and previous MAAS scores are no longer observed, and hence, are not shaded. It is worth noting here that the graph represented in Figure 2.4 corresponds to a Bayesian network of a probabilistic expert system and it is not aimed at modeling the interactions between the variables involved in ICU sedation. The Bayesian network capturing the actual interaction of these variables has a different dependency structure, and hence, its corresponding graph would be different from the graph given in Figure 2.4.

The Bayesian network corresponding to the graph given in Figure 2.4 can be used to determine the proper drug dose for ICU sedation. Specifically, the joint probability distribution $p(u_1, \ldots, u_6, u_1', \ldots, u_6', m, m', b, h, d)$, where $u_1, \ldots, u_5, u_1', \ldots, u_5' \in \{0, 1, 2\}$, $u_6, u_6' \in \{0, \ldots, 10\}$, $m, m' \in \{0, 1, \ldots, 6\}$, $b, h \in \mathbb{R}_+$, and $d \in \{1, 2, \ldots, 12\}$, can be computed using the relationship given in (2.6); namely,

$$p(u_1, \ldots, u_6, u_1', \ldots, u_6', m, m', b, h, d) = \left(\prod_{i=1}^{5} p(u_i') \right) p(u_6' | u_1', \ldots, u_5') p(b) p(h)$$

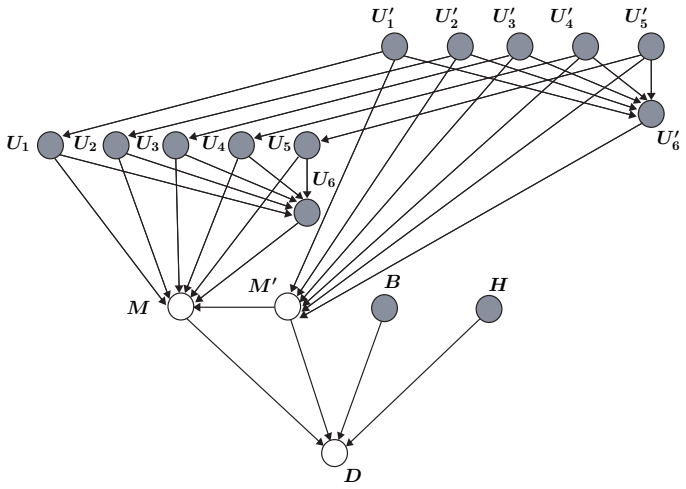

Figure 2.4: A more general graph of a Bayesian network capturing the relationship between the MAAS score and other observable factors; namely, current and previous objective assessments of facial expression U_1, U_1', gross motor movement U_2, U_2', guarding U_3, U_3', heart rate and blood pressure stability U_4, U_4', non-cardiac sympathetic stability U_5, U_5', and non-verbal pain scale U_6, U_6'.

$$\cdot \left(\prod_{i=1}^{5} p(u_i|u_i') \right) p(u_6|u_1, \ldots, u_5, u_6')\, p(m|u_1, \ldots, u_6, m') \cdot p(m'|u_1', \ldots, u_6')\, p(d|m, m', b, h). \quad (2.7)$$

The probability distribution of the drug dose suggested by the Bayesian network is given by

$$\begin{aligned} &p(d|u_1, \ldots, u_6, u_1', \ldots, u_6', b, h) \\ &\qquad = \frac{p(u_1, \ldots, u_6, u_1', \ldots, u_6', b, h, d)}{p(u_1, \ldots, u_6, u_1', \ldots, u_6', b, h)} \\ &\qquad = \frac{\sum_{m=0}^{6} \sum_{m'=0}^{6} p(u_1, \ldots, u_6, u_1', \ldots, u_6', m, m', b, h, d)}{\sum_{m=0}^{6} \sum_{m'=0}^{6} \sum_{d=1}^{12} p(u_1, \ldots, u_6, u_1', \ldots, u_6', m, m', b, h, d)}, \end{aligned} \quad (2.8)$$

where we use marginalization to eliminate m and m', and m, m', and d from the probability density functions in the numerator and denominator of (2.8), respectively. Note that $p(d|m, m', b, h)$, which captures the drug dosing pattern of the medical staff, can be determined through statistical techniques (e.g., maximum likelihood estimates [108]) and clinical data collection. In addition, the probability distributions $p(m|u_1, \ldots, u_6, m')$ and $p(m'|u_1', \ldots, u_6')$ capture the relationship between the

facial expression, gross motor movement, guarding, heart rate and blood pressure stability, noncardiac sympathetic stability, and nonverbal pain scale and the MAAS score, which also requires clinical data collection. The prior probability distributions over $U'_1, \ldots, U'_5$, B, and H denoted by $p(u'_i)$, $i = 1, \ldots, 5$, $p(b)$, and $p(h)$, respectively, as well as the conditional probability distributions $p(u_i|u'_i)$, $i = 1, \ldots, 6$, $p(u'_6|u'_1, \ldots, u'_5)$, and $p(u_6|u_1, \ldots, u_5, u'_6)$ can also be determined by statistical techniques.

Given the probability distribution of the drug dose suggested by the Bayesian network $p(d|u_1, \ldots, u_6, u'_1, \ldots, u'_6, b, h)$, different strategies for choosing the drug dose can be used. One such strategy is the *maximum a posteriori* (MAP) approach [15], where the drug dose corresponding to the mode of the distribution is selected by

$$D_{\text{suggested}} = \text{argmax}_{d \in \{1,\ldots,12\}} p(d|u_1, \ldots, u_6, u'_1, \ldots, u'_6, b, h), \tag{2.9}$$

where $D_{\text{suggested}}$ denotes the drug dose suggested by the Bayesian network.

Finally, note that the Bayesian network can also be used to compute the probability distribution of the drug dose (and hence, the suggested drug dose) when only partial observations are available. In particular, the posterior probability distribution on the drug dose can be computed when the observed variables are a subset of the observed variables in Figure 2.4. Partial observation can result from sensor failure, where a particular state of the patient is unavailable at the time of a decision. For example, if only the blood pressure, heart rate, facial expression, and gross motor movement data is available, $p(d|u_1, u_2, u'_1, u'_2, b, h)$ gives the probability distribution for the drug dose based on these partial observations. The graph of the Bayesian network for this case is given in Figure 2.5.

The same probabilistic methodology can be used to account for the secondary action in the Bayesian network. In this case, appropriate random variables have to

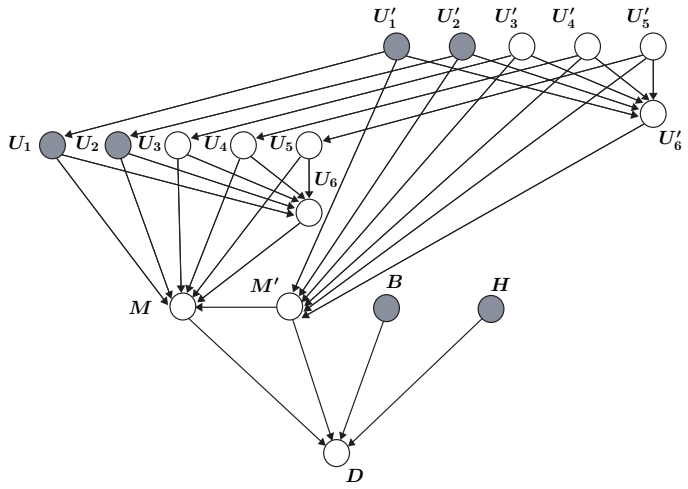

Figure 2.5: A graph of a Bayesian network where only the current gross motor movement U_1, previous gross motor movement U_2, blood pressure B, and heart rate H are observable.

be defined and the graph given in Figure 2.4 should be modified accordingly. Alternatively, a hybrid probabilistic-deterministic expert system can be defined, where the primary actions given by the first part of the conclusions in Table 2.1 are described by a Bayesian network and the secondary action is given by Table 2.2 and the flowchart given in Figure 2.2. More specifically, define the switching random variable S, where range$(S) = \{0, 1\}$. The random variable S acts as a switch, where $S = 0$ denotes that the HDESC is off-line and $S = 1$ denotes that the HDESC is activated. The activation could be probability-based where the HDCES is activated if $P(S = 1|u_1, \ldots, u_6, u_1', \ldots, u_6', b, h) > T$, where $0 < T < 1$, $P(\cdot|\cdot)$ is the conditional probability operator, and T is a threshold value. Note that based on Table 2.1, the activation of the HDCES depends on the current and previous MAAS scores, blood pressure, heart rate, and the required drug dose for ICU sedation. The graph of the hybrid Bayesian network is shown in Figure 2.6.

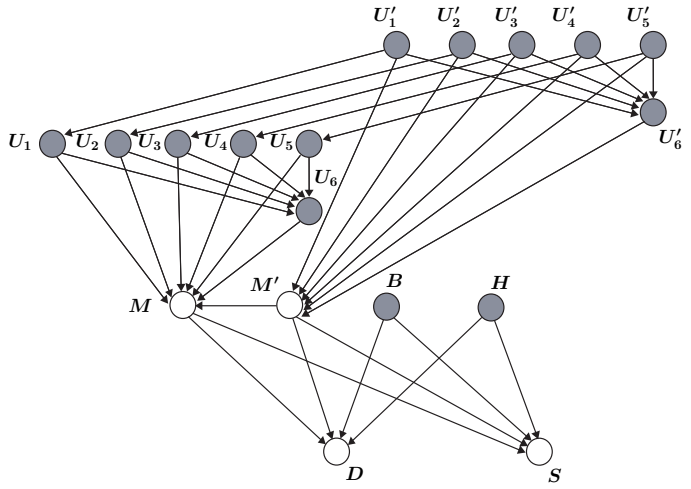

Figure 2.6: A graph of a hybrid probabilistic-deterministic Bayesian network where the random variable S controls the activation of the hemodynamic control expert system.

2.8. A Probabilistic Alarm Algorithm for Critical Care Monitoring

A potentially key application of the probabilistic expert system developed in Section 2.7 is its applicability to clinical decision support, critical care monitoring, and lifesaving interventions. A clinical decision support system is a computer program that can directly provide the medical staff with assessments and recommendations in the clinical decision making process [75]. A clinical decision support system can be coupled to a closed-loop control system to provide a hierarchical hybrid control architecture characterized by continuous-time control algorithms at the lower-level units and logical decision-making units at the higher-level of the hierarchy. In particular, a hybrid controller would involve the clinician evaluating the patient through a decision support system and an autonomous closed-loop controller adjusting the desired regimen to maintain sedation at a desired level. This controller architecture allows for the expert system to directly aid in clinical decision making as well as critical care monitoring and lifesaving interventions.

In this section, we use the framework presented in Section 2.7 to design an alarm algorithm for agitation detection in ICU patients. An alarm system refers to an automatic warning system that constantly monitors a specific state of the patient and notifies the medical staff in case of an abnormality [78]. An agitation detection alarm system can reduce the medical staff's workload as well as safeguard against life-threatening situations in the ICU.

To design an alarm algorithm for agitation detection, let the patients's facial expression, gross motor movement, guarding, heart rate and blood pressure stability, noncardiac sympathetic stability, nonverbal pain scale, and patient agitation be given by the random variables $U_1, \ldots, U_6$, and A, respectively, where range$(U_i) = \{0, 1, 2\}$, $i = 1, \ldots, 5$, range$(U_6) = \{0, \ldots, 10\}$, and range$(A) = \{0, 1\}$. Here, $A = 0$ and $A = 1$ denote, respectively, absence of agitation corresponding to an MAAS score of 0–3 and presence of restlessness and agitation corresponding to an MAAS score of 4–6. The graph of the Bayesian network for the alarm algorithm is given by Figure 2.7. Note that the random variables $U_1, \ldots, U_6$ are considered as observed variables, and hence, are shaded.

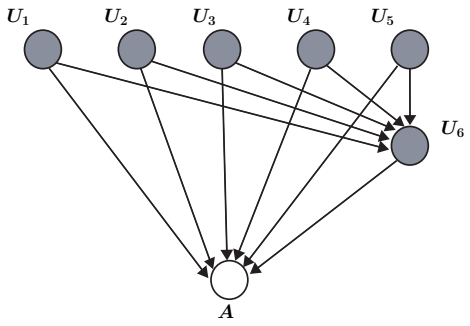

Figure 2.7: A graph of a Bayesian network of an alarm algorithm for critical care monitoring.

The joint probability distribution for this network is given by

$$p(u_1, \ldots, u_6, a) = p(u_1) \ldots p(u_5) p(u_6|u_1, \ldots, u_5) p(a|u_1, \ldots, u_6), \tag{2.10}$$

where $u_i \in \{0, 1, 2\}$, $i = 1, \ldots, 5$, $u_6 \in \{0, \ldots, 10\}$, and $a \in \{0, 1\}$. In addition, the posterior probability of the patient's agitated state is given by

$$p(a|u_1, \ldots, u_6) = \frac{p(u_1, \ldots, u_6, a)}{p(u_1, \ldots, u_6)} = \frac{p(u_1, \ldots, u_6, a)}{\sum_{a=0}^{1} p(u_1, \ldots, u_6, a)}. \tag{2.11}$$

The patient's agitation state is based on the observation of facial expression, gross motor movement, guarding, heart rate and blood pressure stability, noncardiac sympathetic stability, and nonverbal pain scale and is given by

$$A_{\text{predicted}} = \text{argmax}_{a \in \{0,1\}} p(a|u_1, \ldots, u_6), \tag{2.12}$$

where we use the MAP approach to select the agitation state with the highest probability.

To elucidate the efficacy of our proposed approach we apply our framework to a retrospective study involving recorded sedation and agitation data for 366 patients admitted to the ICU in Northeast Georgia Medical Center, Gainesville, GA, over the period of May 6, 2009 to April 27, 2010. The patient's age ranged from 18 to 90 years. The length of stay in the ICU ranged from 1 to 93 days. In addition, the available data set included 15,052 measurements of facial expressions, gross motor movement, guarding, heart rate and blood pressure stability, noncardiac sympathetic stability, nonverbal pain scale, and agitation state. We used the hold-out method [15] for validation of the alarm algorithm, where 12,000 measurements of the random variables $U_1, \ldots, U_6$ and A were used to train the Bayesian network. The algorithm was tested on the remaining 3,052 measurements of $U_1, \ldots, U_6$.

We used the MATLAB® version R2008a and the Bayesian Network MATLAB® Toolbox [109] to compute the posterior probability distributions of the patient's agitation state. In the training stage, we used the maximum likelihood estimates approach to estimate the probability distributions in (2.10) and used a *uniform Dirichlet prior* to avoid zero conditional probabilities for cases not present in the training data

Table 2.4: Comparison of the human-assessed patient agitation with the predicted agitation state given by the alarm algorithm.

	Patient Not Agitated	Patient Agitated
Alarm On	48	206
Alarm Off	2628	125

set [108]. In addition, in the testing stage, we used the *junction-tree inference algorithm* [15]. Table 2.4 gives the predicted agitation state $A_{\text{predicted}}$ given by (2.12) as compared to the human-assessed patient agitation. Based on the results, for any given alarm, that is, when the algorithm predicts the presence of patient agitation, the probability that the patient is not agitated (false positive) is 18.9% with a 95% confidence interval of 14.4% to 24.4%. In addition, when no agitation is reported by the algorithm the probability that the patient is experiencing agitation (false negative) is 4.5% with a 95% confidence interval of 3.8% to 5.4%. The confidence interval is calculated based on the framework presented in [114]. In 45 cases out of a total of 3,052 cases the algorithm was undecided; that is, the posterior probability of the patient's agitation state was uniform.

2.9. ICU Sedation Control

In this section, we present an illustrative numerical example of a probabilistic expert system for ICU sedation control based on the discussion given in Section 2.7. For illustrative purposes, however, we consider a simplified probabilistic expert system to control ICU sedation. Specifically, let the random variables U_1 and U_2 represent the current objective assessment of the facial expression and gross motor movement, respectively, and let U_1' and U_2' represent the previous objective assessment of the facial expression and gross motor movement, respectively. Furthermore, let the patient's current MAAS score, previous MAAS score, and required drug dose for seda-

tion be given by the random variables M, M', and D, respectively. Moreover, let range(U_1) = range(U_1') = $\{0, 1\}$, where 0 and 1 denote a relaxed face and grimacing and moaning face, respectively, let range(U_2) = range(U_2') = $\{0, 1, 2\}$, where 0, 1, and 2 denote lying quietly, cautious movement, and restless withdrawal, respectively, let range(M) = range(M') = $\{1, 2, 3\}$, where 1, 2, and 3 denote MAAS scores 0, 3, and 6, respectively, and let range(D) = $\{1, 2, 3\}$, where 1, 2, and 3 denote decrease in propofol infusion, no action, and increase in propofol infusion, respectively. It is important to note here that even though a simplified set of motor movements and facial expressions, MAAS scores, and drug dosing requirements are used to develop this expert system, this is made for simplicity of exposition and is not clinically unrealistic.

The graph of the Bayesian network which includes these random variables is given in Figure 2.8. The conditional probability distributions are given by Table 2.5. Here, we assume a uniform prior probability distributions for the random variables U_1, U_2, U_1', and U_2'. Note that U_1, U_2, U_1', and U_2' are observed random variables, and hence, their corresponding nodes in the graph given by Figure 2.8 are shaded. The conditional probability distributions can be determined through statistical techniques and clinical data collection which correspond to a classical interpretation of probability. Alternatively, the conditional probabilities can also reflect the expertise of a human expert which corresponds to a Bayesian interpretation of probability. As noted above, this interpretation of probabilities can be used to quantify uncertainty.

We used the MATLAB® version R2008a and the Bayesian Network MATLAB® Toolbox [109] to compute the posterior probability distributions of the required action, where we used the *junction-tree inference algorithm* [15]. The posterior probability distributions of the required action D for the case where $U_1 = 0$, $U_2 = 0$, $U_1' = 1$, $U_2' = 1$, and $U_1 = 1$, $U_2 = 2$, $U_1' = 0$, $U_2' = 1$ are given in Figures 2.9 and 2.10,

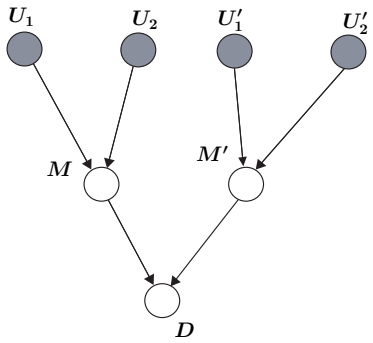

Figure 2.8: The graph of a Bayesian network capturing the relationships between the current MAAS score (M), previous MAAS score (M'), current and previous facial expressions (U_1, U_1'), current and previous gross motor movement (U_2, U_2'), and required drug dose (D).

respectively. If a maximum a posteriori (MAP) strategy is used, then the action with the highest posterior probability is selected. This corresponds to "decrease propofol infusion" and "increase propofol infusion" for the examples given in Figures 2.9 and 2.10, respectively. The posterior probability distributions of the required action in the case where only the random variable corresponding to the gross motor movement is observed (i.e., $U_2 = 0$ and $U_2' = 1$) is given in Figure 2.11.

Comparing Figure 2.9 with Figure 2.11 it follows that the posterior probability distribution of the required action D changes due to increase in uncertainty. The Bayesian network can also be used to find the posterior probability distribution for the hidden variables in the graph. The posterior probability distribution for the current MAAS score M for the case $U_1 = 1$, $U_2 = 2$, $U_1' = 0$, and $U_2' = 1$ is given in Figure 2.12. Finally, note that the posterior probability distribution reflects the uncertainty associated with each value the random variable assumes. For example, using Figure 2.12, the posterior probability of M reflects the fact that there is a high likelihood that the patient is highly agitated, which justifies the required action "increase propofol infusion."

Table 2.5: Conditional probability distributions for M, M', and D.

U_1, U_1'	U_2, U_2'	M, M'	P
0	0	1	0.8
0	0	2	0.15
0	0	3	0.05
1	0	1	0.3
1	0	2	0.35
1	0	3	0.35
0	1	1	0.15
0	1	2	0.8
0	1	3	0.05
1	1	1	0.1
1	1	2	0.4
1	1	3	0.5
0	2	1	0.05
0	2	2	0.35
0	2	3	0.6
1	2	1	0.05
1	2	2	0.15
1	2	3	0.8

M	M'	D	P
1	1	1	0.8
1	1	2	0.19
1	1	3	0.01
1	2	1	0.9
1	2	2	0.09
1	2	3	0.01
1	3	1	0.95
1	3	2	0.04
1	3	3	0.01
2	1	1	0.25
2	1	2	0.7
2	1	3	0.05
2	2	1	0.05
2	2	2	0.9
2	2	3	0.05
2	3	1	0.5
2	3	2	0.45
2	3	3	0.05
3	1	1	0.01
3	1	2	0.09
3	1	3	0.9
3	2	1	0.01
3	2	2	0.19
3	2	3	0.8
3	3	1	0.01
3	3	2	0.04
3	3	3	0.95

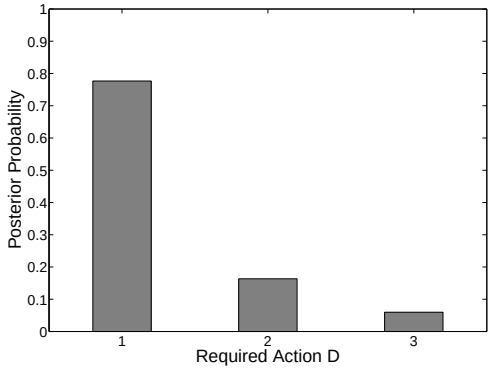

Figure 2.9: Posterior probability distribution of D for the case $U_1 = 0$, $U_2 = 0$, $U_1' = 1$, and $U_2' = 1$.

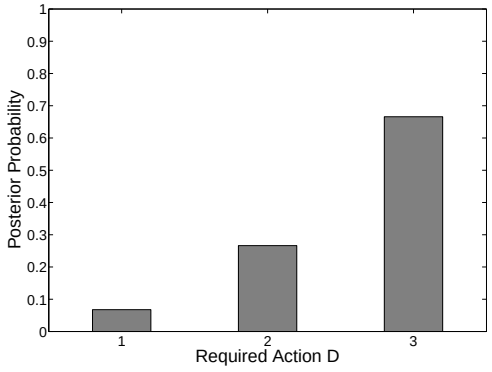

Figure 2.10: Posterior probability distribution of D for the case $U_1 = 1$, $U_2 = 2$, $U_1' = 0$, and $U_2' = 1$.

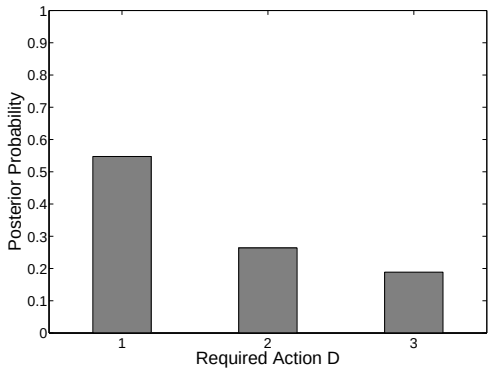

Figure 2.11: Posterior probability distribution of D for the case $U_2 = 0$ and $U_2' = 1$.

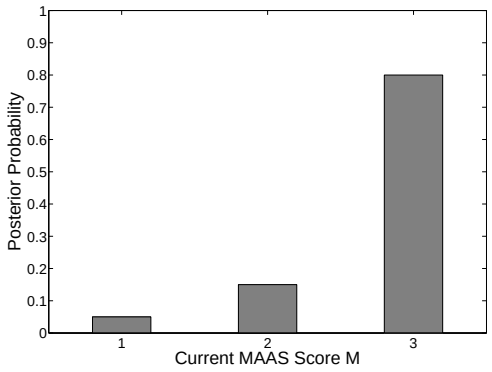

Figure 2.12: Posterior probability distribution of M for the case $U_1 = 1$, $U_2 = 2$, $U_1' = 0$, and $U_2' = 1$.

Chapter 3

Relevance Vector Machine Learning for Neonate Pain Intensity Assessment Using Digital Imaging

3.1. Introduction

Individuals in pain manifest their condition through "pain behavior" [94, 120], which includes facial expressions. In [120], the significance of a facial expression as an indicator of pain is discussed and advances in pain assessment using facial expressions are reviewed. Clinicians regard the patient's facial expression as a valid indicator for pain and pain intensity [38]. Hence, correct interpretation of the facial expressions of the patient and its correlation with pain is a fundamental step in designing an automated pain assessment management system. Of course, other pain behaviors including head movement and the movement of other body parts, along with physiological indicators of pain, such as heart rate, blood pressure, and respiratory rate responses should also be included in such a system.

Depending on the patient group (e.g., neonates, children, adults, etc.) pain assessment criteria have been developed, and indicators of pain in each group might be different. For example, while the behavioral pain scale for adults focuses on facial expressions, upper limbs, and compliance with ventilation [118], the face, legs, activity, cry, and consolability (FLACC) [107] behavioral pain scale focuses on slightly

different set of indicators for postoperative young children. Similarly, the premature infant pain profile (PIPP) [134] considers a special set of pain indicators including physiological and behavioral indicators for pain assessment in premature infants.

Infants are unable to directly report their level of pain, and hence, medical staff are responsible for pain assessment for neonates. Pain and distress behaviors in neonates include facial expression, cry, and body movement, and a series of methods have been suggested to objectively assess pain in neonates based on the aforementioned behaviors [107, 120, 134]. In this chapter, we focus on the problem of pain assessment in infants using facial expressions.

Although there is a vast potential for using computer vision for agitation and pain assessment, there are very few articles in the computer vision literature addressing this issue. The authors in [17] have used computer vision for pain assessment in demented elderly patients. In [6], an agitation assessment scheme is proposed for patients in the ICU. The approach of [6] is based on the hypothesis that facial grimacing induced by pain results in additional "wrinkles" (equivalent to edges in the processed image) on the face of the patient, and this is the only factor they use in assessing pain. Although this approach is computationally inexpensive and especially appealing for a real-time decision support system, it can be limiting since it does not account for other facial actions (e.g., smiling, crying, etc.), which may not necessarily correspond to pain. The authors in [20–23] use various face classification techniques including support vector machines (SVM) and neural networks (NN) to classify facial expressions in neonates into "pain" and "non-pain" classes. Such classification techniques were shown to have reasonable accuracy.

In this chapter, we extend the classification technique addressed in [20–23] to distinguish pain from non-pain in neonates as well as assess their pain intensity using a relevance vector machine (RVM) classification technique [139]. The RVM classifi-

cation technique is a Bayesian extension of SVM which achieves comparable performance to SVM while providing posterior probabilities for class memberships and a sparser model. In a Bayesian interpretation of probability, as opposed to the classical interpretation, the probability of an event is an indication of the *uncertainty* associated with the event rather than its *frequency* [15]. If data classes represent "pure" facial expressions, that is, extreme expressions that an observer can identify with a high degree of confidence, the posterior probability of the membership of some intermediate facial expression to a class can provide an estimate of the intensity of such an expression. This, along with other pain behaviors, can be translated into one of the scoring systems currently being used for assessing pain (e.g., FLACC or PIPP).

The contents of the chapter are as follows. In Sections 3.2 and 3.3, we review the SVM and RVM classification techniques for pain recognition using facial expressions. Then, in Section 3.4, we present the results of these classification techniques applied to the Infant Classification of Pain Expression (COPE) database [22]. The pain intensity assessment given by the computer classifier shows a strong correlation with the pain intensity assessed by expert and non-expert human examiners.

3.2. Support Vector Machines

As we see in Section 3.4, the problem of pain and pain intensity assessment using facial images involves a standard problem in machine learning called *data classification* [15]. Given a series of input variables $x_1, x_2, \ldots, x_N$ in $\mathbb{R}^D$ and their corresponding class labels $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_p$, where $p \leq N$, the data classification problem involves assigning the correct class label to a new input variable x. Kernel-based methods are typically used for data classification and regression [15]. A key limitation of many kernel-based learning algorithms is the computational intensity involved in the training, prediction, and decision making stages of the algorithm. This is due

to the fact that the kernel function, which adds a dimension to the data in order to obtain an optimal classification, has to be computed for all pairs of data points. In sparse kernel machine algorithms, however, only a subset of the training data is used, providing a sparse solution. Sparse kernel machines are faster in the training, and the prediction and decision making stages. In this chapter, we consider two sparse kernel-based classification algorithms, namely, support vector machines and relevance vector machines.

Support vector machines (SVM) [144] involve sparse kernel algorithms used in classification and regression problems, and have their origin in statistical learning theory. Here, we consider the classification problem involving two data classes, namely, $\mathcal{C}_1$ and $\mathcal{C}_2$. The framework can be generalized to a multi-class label problem using a similar approach as outlined below [15]. Let the *training set* be given by $\{x_1, x_2, \ldots, x_N\}$, with *target values* given by $z_1, z_2, \ldots, z_N$, respectively, where $x_n \in \mathbb{R}^D$ and $z_n \in \{-1, 1\}$, $n = 1, \ldots, N$, and with $x_n \in \mathcal{C}_1$ if $z_n = -1$, and $x_n \in \mathcal{C}_2$ if $z_n = 1$. To classify a new data point $x \in \mathbb{R}^D$, define the *classifier function*

$$y(x) \triangleq w^{\mathrm{T}}\phi(x) + b, \tag{3.1}$$

where $\phi : \mathbb{R}^D \to \mathbb{R}^M$ is a continuous fixed feature-space transformation, $w \in \mathbb{R}^M$ is a weight vector, and $b \in \mathbb{R}$ is a bias parameter. The sign of the classifier function $y(x)$ determines the class of x. More specifically, for a new input variable x, the target value is given by $z = \operatorname{sgn}(y(x))$, where $\operatorname{sgn} y \triangleq \frac{y}{|y|}$, $y \neq 0$, and $\operatorname{sgn}(0) \triangleq 0$.

Next, assume that the training set is linearly separable in the feature space $\mathbb{R}^M$, that is, there exist a weight vector $w \in \mathbb{R}^M$ and a bias parameter $b \in \mathbb{R}$ such that $y(x_n) > 0$ for $x_n \in \mathbb{R}^D$ and $z_n = 1$, and $y(x_n) < 0$ for $x_n \in \mathbb{R}^D$ and $z_n = -1$; or, equivalently, $z_n y(x_n) > 0$ for all $x_n \in \mathbb{R}^D$ and $z_n \in \{-1, 1\}$. Later we will relax the linear separability assumption and consider the more general case of overlapping

classes.

Note that the classifier function $y(\cdot)$ separates the feature space $\mathbb{R}^M$ into two disjoint regions characterized by $y(x) > 0$ and $y(x) < 0$ for $x \in \mathbb{R}^D$. The affine hyperplane separating the two disjoint regions, namely $y(x) = 0$, is called the *decision boundary* and is denoted by $\mathcal{D}$. Note that $\phi(\cdot)$ can be a nonlinear transformation, which would correspond to a nonlinear decision boundary in the original input space $\mathbb{R}^D$. The minimum distance between the training set and the decision boundary $\mathcal{D}$ is called the *margin*. The distance of a point $x_n \in \mathbb{R}^D$ to the decision boundary $\mathcal{D}$ is given by

$$\text{dist}(\phi(x_n), \mathcal{D}) = \frac{|y(x_n)|}{\|w\|} = \frac{z_n y(x_n)}{\|w\|}, \tag{3.2}$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^M$ and $\text{dist}(x, \mathcal{D}) \triangleq \inf_{s\in\mathcal{D}} \|x - s\|$. Hence, the margin is given by

$$\min_{n\in\{1,\ldots,N\}} \text{dist}(\phi(x_n), \mathcal{D}) = \min_{n\in\{1,\ldots,N\}} \frac{z_n y(x_n)}{\|w\|}. \tag{3.3}$$

As in all classification methods, the goal of the SVM algorithm is to classify a new input variable $x \in \mathbb{R}^D$ based on the information provided by the training set and the target values. The SVM framework addresses this problem by choosing the decision boundary in such a way so that the margin is maximized. The following problem presents the SVM algorithm as an optimization problem.

Maximum Margin Classification Problem. Consider the training set given by $\{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$ and let the classifier function $y : \mathbb{R}^D \to \mathbb{R}$ be given by (3.1). Find the weight vector $w \in \mathbb{R}^M$ and the bias parameter $b \in \mathbb{R}$ such that (3.3) is maximized.

Theorem 3.1. $w^* \in \mathbb{R}^M$ and $b^* \in \mathbb{R}$ solve the Maximum Margin Classification

Problem if and only if w^* and b^* are the solutions to the optimization problem

$$\max_{w\in\mathbb{R}^M,\, b\in\mathbb{R}} \left\{ \frac{1}{\|w\|} \min_{n\in\{1,\ldots,N\}} \left[z_n \left(w^{\mathrm{T}}\phi(x_n) + b \right) \right] \right\}. \tag{3.4}$$

Proof. The proof is a direct consequence of the definition of a margin given by (3.3). □

The solution to the non-convex optimization problem (3.4) is not unique. To see this, note that scaling the weight vector $w \in \mathbb{R}^M$ and the bias parameter $b \in \mathbb{R}$ by a positive scalar does not change the value of the function to be maximized. The following theorem presents an alternative characterization to the Maximum Margin Classification Problem.

Theorem 3.2. $w^* \in \mathbb{R}^M$ and $b^* \in \mathbb{R}$ solve the Maximum Margin Classification Problem if and only if w^* and b^* are the solutions to the optimization problem

$$\min_{w\in\mathbb{R}^M,\, b\in\mathbb{R}} \frac{1}{2}\|w\|^2 \tag{3.5}$$

subject to

$$z_n \left(w^{\mathrm{T}}\phi(x_n) + b \right) \geq 1, \tag{3.6}$$

where $x_n \in \mathbb{R}^D$, $z_n \in \{-1, 1\}$, and $n = 1, \ldots, N$.

Proof. Since rescaling the weight vector $w \in \mathbb{R}^M$ and the bias parameter $b \in \mathbb{R}$ in (3.4) by a positive scalar does not change the value of the function to be maximized, the optimization problem (3.4) has a continuum of solutions corresponding to the same optimal value. Hence, introducing the new constraint

$$z_{n^*} \left(w^{\mathrm{T}}\phi(x_{n^*}) + b \right) = 1, \tag{3.7}$$

where $n^* = \arg\min_{n\in\{1,\ldots,N\}} \mathrm{dist}(\phi(x_n), \mathcal{D})$, does not change the optimal value of the optimization problem (3.4). Thus, the inequality constraint (3.6) holds for all

$x_n \in \mathbb{R}^D$, $z_n \in \{-1, 1\}$, and $n = 1, \ldots, N$. The proof now follows by noting that the optimization problem (3.4) subject to (3.7) is equivalent to the optimization problem (3.5) subject to (3.6). □

The constrained optimization problem given by (3.5) and (3.6) is convex and can be solved using Lagrange multiplier methods. Specifically, introducing the Lagrange multipliers $\lambda_n \in \mathbb{R}$, $n = 1, \ldots, N$, and forming the Lagrangian

$$\mathcal{L}(w, b, \lambda) = \frac{1}{2}\|w\|^2 - \sum_{n=1}^{N} \lambda_n \left[z_n \left(w^{\mathrm{T}} \phi(x_n) + b \right) - 1 \right], \tag{3.8}$$

where $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]^{\mathrm{T}}$, it follows from the first-order necessary conditions for optimality that

$$w = \sum_{n=1}^{N} \lambda_n z_n \phi(x_n), \tag{3.9}$$

$$0 = \sum_{n=1}^{N} \lambda_n z_n. \tag{3.10}$$

Note that (3.9) and (3.10) can be used to eliminate w and b from the Lagrangian (3.8) leading to a *dual representation* of the optimization problem (3.4). Namely,

$$\max_{\lambda \in \mathbb{R}^N} \tilde{\mathcal{L}}(\lambda) \tag{3.11}$$

subject to

$$\lambda_n \geq 0, \quad n = 1, \ldots, N, \tag{3.12}$$

$$\sum_{n=1}^{N} \lambda_n z_n = 0, \tag{3.13}$$

where

$$\tilde{\mathcal{L}}(\lambda) \triangleq \sum_{n=1}^{N} \lambda_n - \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \lambda_n \lambda_m z_n z_m k(x_n, x_m) \tag{3.14}$$

and

$$k(x, x') = \phi^{\mathrm{T}}(x) \phi(x') \tag{3.15}$$

is the kernel function. Here, we introduced an alternative formulation of the optimization problem (3.4) in terms of the kernel function $k : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$, which allows us to avoid working explicitly in the feature space. Note that the classifier function (3.1) can be rewritten using the kernel function as

$$y(x) = \sum_{n=1}^{N} \lambda_n z_n k(x, x_n) + b. \tag{3.16}$$

The Kuhn-Tucker (KT) necessary conditions for optimality for the constrained optimization problem (3.11)–(3.13) are given by

$$\lambda_n \geq 0, \tag{3.17}$$

$$z_n y(x_n) - 1 \geq 0, \tag{3.18}$$

$$\lambda_n \left(z_n y(x_n) - 1 \right) = 0, \tag{3.19}$$

where $n = 1, \ldots, N$. Now, it follows from (3.19) that either $\lambda_n = 0$ or $z_n y(x_n) = 1$. The input variables $x_n \in \mathbb{R}^D$, $n = 1, \ldots, N$, for which the corresponding Lagrange multiplier $\lambda_n \in \mathbb{R}$ vanishes do not contribute to the classifier function (3.16) and, hence, can be omitted. The remaining input variables are called *support vectors* and, by definition, lie on the maximum margin affine hyperplanes ${w^*}^{\mathrm{T}} \phi(x_n) + b^* = \pm 1$, $n = 1, \ldots, N$. Hence, only the support vectors play a role in the classification of the new input variables and the rest of the training set can be discarded.

Next, we consider the case of overlapping classes. For this case, the SVM algorithm considered above identifies the decision boundary so that the training set is separated into two data classes with no input variables being misclassified. This results in poor class assignments for new input variables. The SVM algorithm, however, can be modified by allowing input variables in the training set to lie on the "wrong side" of the margin boundary and penalizing such constraint violations. Specifically, for every input variable $x_n \in \mathbb{R}^D$, $n = 1, \ldots, N$, define the *slack variable* $\xi_n \geq 0$ such

that $\xi_n = 0$ if (3.6) is satisfied, that is, for $n \in \{1, \ldots, N\}$, x_n is on or inside the correct margin boundary, and $\xi_n = |z_n - y(x_n)|$ otherwise.

The modified SVM algorithm is given by the following optimization problem

$$\min_{w \in \mathbb{R}^M,\, b \in \mathbb{R},\, \xi \in \mathbb{R}^N} C \sum_{n=1}^{N} \xi_n + \frac{1}{2}\|w\|^2 \tag{3.20}$$

subject to

$$z_n y(x_n) \geq 1 - \xi_n, \quad n = 1, \ldots N, \tag{3.21}$$

$$\xi_n \geq 0, \tag{3.22}$$

where $\xi = [\xi_1, \xi_2, \ldots, \xi_N]^{\mathrm{T}}$ and $C > 0$ is a *complexity parameter* controlling the trade-off between the margin and the slack variable penalty. It can be shown that if $\xi_n = 0$, then (3.21) reduces to (3.6) and the corresponding input variable $x_n \in \mathbb{R}^D$ will be correctly classified. Moreover, if $0 < \xi_n \leq 1$, then the input variable $x_n \in \mathbb{R}^D$ is correctly classified while lying inside the margin boundary, whereas if $\xi_n > 1$, then the input variable is misclassified.

Lagrange multiplier methods can be used to solve the optimization problem (3.20)–(3.22) by introducing the Lagrange multipliers $\lambda_n \in \mathbb{R}$ and $\mu_n \in \mathbb{R}$, $n = 1, \ldots, N$, corresponding to the constraints (3.21) and (3.22), respectively. In this case, the Lagrangian is given by

$$\mathcal{L}(w, b, \lambda, \xi) = \frac{1}{2}\|w\|^2 + C \sum_{n=1}^{N} \xi_n - \sum_{n=1}^{N} \lambda_n \left(z_n y(x_n) - 1 + \xi_n\right) - \sum_{n=1}^{N} \mu_n \xi_n, \tag{3.23}$$

where $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_N]^{\mathrm{T}}$. Now, it follows from the first-order necessary conditions for optimality that

$$w = \sum_{n=1}^{N} \lambda_n z_n \phi(x_n), \tag{3.24}$$

$$0 = \sum_{n=1}^{N} \lambda_n z_n, \tag{3.25}$$

$$\lambda_n = C - \mu_n, \quad n = 1, \ldots, N, \tag{3.26}$$

and the KT necessary conditions give

$$\lambda_n \geq 0, \tag{3.27}$$

$$z_n y(x_n) - 1 + \xi_n \geq 0, \tag{3.28}$$

$$z_n \left(z_n y(x_n) - 1 + \xi_n \right) = 0, \tag{3.29}$$

$$\mu_n \geq 0, \tag{3.30}$$

$$\xi_n \geq 0, \tag{3.31}$$

$$\mu_n \xi_n = 0, \tag{3.32}$$

where $n = 1, \ldots, N$. The dual representation of the optimization problem (3.20) subject to (3.21) and (3.22) is given by

$$\min_{\lambda \in \mathbb{R}^N} \tilde{\mathcal{L}}(\lambda) \tag{3.33}$$

subject to

$$0 \leq \lambda_n \leq C, \quad n = 1, \ldots, N, \tag{3.34}$$

$$\sum_{n=1}^{N} \lambda_n z_n = 0, \tag{3.35}$$

where $\tilde{\mathcal{L}}(\lambda)$ is given by (3.14), the kernel function is given by (3.15), and where we have used (3.24)–(3.32).

3.3. Sparse Bayesian Learning

The SVM framework is a powerful classifier but has a number of limitations. A key deficiency of the approach is the fact that the output of the SVM is the binary classification decision and not the class membership posterior probability. As will be discussed in Section 3.4, methods which possess an inherent Bayesian structure are more powerful and can provide more information. Such methods not only classify a new input variable, but can also provide a degree of uncertainty (in terms of posterior

probabilities) for such a classification. The relevance vector machine (RVM) [139], which is a special case of the sparse Bayesian learning algorithm, can be regarded as the Bayesian extension of the SVM approach.

In this section, we consider a classification problem involving two data classes, namely $\mathcal{C}_1$ and $\mathcal{C}_2$, using the sparse Bayesian learning approach. The framework can be generalized to a multi-class classification problem using a similar approach as outlined below [139]. Consider the Laplace approximation method [15] involving the random variable $v \in \mathbb{R}^M$ with associated probability density function given by $p : \mathbb{R}^M \to \mathbb{R}$. Assume that $p(v) = \frac{f(v)}{V}$, where $f : \mathbb{R}^M \to \mathbb{R}$ is a function defined on $v \in \mathbb{R}^M$ and $V = \int_{\mathbb{R}^M} f(v)\mathrm{d}v$ is the normalization coefficient. The probability density function $p(v)$ is approximated by a multivariate Gaussian (normal) distribution $\mathcal{N}(v; v_0, \Sigma)$ with mean $v_0 \in \mathbb{R}^M$ and covariance matrix $\Sigma \in \mathbb{R}^{M\times M}$, where $v_0 = \arg\max_{v\in\mathbb{R}^M} p(v)$ and $\Sigma = -\frac{\partial^2}{\partial v^2} \ln f(v)|_{v=v_0}$. The normalization coefficient V can be approximated by [15]

$$V \simeq f(v_0)\frac{(2\pi)^{M/2}}{(\det \Sigma)^{\frac{1}{2}}}, \tag{3.36}$$

where $\det(\cdot)$ denotes the determinant operator.

Next, let the training set be given by $\{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$, with target values given by $z_1, z_2, \ldots, z_N$, respectively, where $x_n \in \mathbb{R}^D$ and $z_n \in \{0, 1\}$, $n = 1, \ldots, N$, and with $x_n \in \mathcal{C}_1$ if $z_n = 1$, and $x_n \in \mathcal{C}_2$ if $z_n = 0$. For a new input variable $x \in \mathbb{R}^D$, we predict the associated class membership posterior probability distribution $p(\mathcal{C}_k|x, X, Z)$, $k = 1, 2$, where $p(\mathcal{C}_k|x, X, Z)$ is the class membership conditional probability of the data class $\mathcal{C}_k$ given $x \in \mathbb{R}^D$, $X = [x_1,\, x_2,\, \ldots,\, x_N]$, and $Z = [z_1,\, z_2,\, \ldots,\, z_N]^{\mathrm{T}}$. Note that, in contrast to the SVM approach, the sparse Bayesian learning method separates the *prediction* stage (i.e., finding the posterior class membership probabilities for the new input variable x) from the *decision making* stage (i.e., assigning the new

input variable x to the appropriate class). This separation is particularly useful when dealing with asymmetric classification costs, where misclassification of input variables belonging to a certain class is more costly [139]. For example, for the problem involving the classification of facial images of patients to pain and non-pain classes discussed in Section 3.4, the cost of misclassification of a patient in pain to the non-pain class (*false negative*) is higher than that of a patient with no pain to the pain class (*false positive*). One of the key advantages of the sparse Bayesian learning approach is its ability to deal with such asymmetric costs.

Define the classifier function

$$y(x) \triangleq w^{\mathrm{T}}\phi(x), \tag{3.37}$$

where $\phi : \mathbb{R}^D \rightarrow \mathbb{R}^M$ is a continuous feature-space transformation and $w = [w_1, w_2, \ldots, w_M]^{\mathrm{T}} \in \mathbb{R}^M$ is a weight vector. Note that the RVM algorithm is a special case of the sparse Bayesian learning method. Specifically, in the RVM, $w^{\mathrm{T}}\phi(x)$ in (3.37) has the special form (similar to the SVM algorithm) given by $\sum_{n=1}^{N} w_n k(x, x_n) + b$, where $k(\cdot, \cdot)$ is the kernel function. In the sequel, we consider the general formulation (3.37).

Following standard statistical practice, we assume that the posterior probability of the target value of an input variable corresponding to the class $\mathcal{C}_1$ is given by $p(z = 1|x_n, w) = \sigma(y(x_n))$, $n = 1, \ldots, N$, where $\sigma(\cdot)$ is the logistic sigmoidal function defined by $\sigma(s) \triangleq \frac{1}{1+\mathrm{e}^{-s}}$ [15, 139]. Note that, since there are only two classes, $p(z = 0|x_n, w) = 1 - \sigma(y(x_n))$. Assuming that the input variables x_n, $n = 1, \ldots, N$, are independent, the *likelihood function* is given by

$$p(Z|X, w) = \prod_{n=1}^{N} p(z|x_n, w) = \prod_{n=1}^{N} \sigma\left(y(x_n)\right)^{z_n} \left(1 - \sigma\left(y(x_n)\right)\right)^{1-z_n}. \tag{3.38}$$

Each weight parameter w_n, $n = 1, \ldots, M$, in (3.37) is assumed to have a zero-mean

Gaussian distribution, and hence, the weight prior distribution is given by

$$p(w|\alpha) = \prod_{n=1}^{M} \mathcal{N}(w_n; 0, \alpha_n^{-1}), \tag{3.39}$$

where α_n, $n = 1, \ldots, M$, is the *precision* (inverse of the variance of the Gaussian distribution) corresponding to w_n and $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_M]^{\mathrm{T}} \in \mathbb{R}^M$. The parameters α_n, $n = 1, \ldots, M$, in the prior distribution (3.39) are called the *hyperparameters*. Note that, in contrast to other Bayesian classifiers, each weight parameter w_n, $n = 1, \ldots, M$, has a separate hyperparameter α_n.

Given a new input variable $x \in \mathbb{R}^D$, the corresponding target value $z \in \{0, 1\}$ can be predicted using the *predictive distribution* $p(z|x, X, Z)$. The predictive distribution is given by

$$p(z|x, X, Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(z|x, X, Z, w, \alpha) p(w|x, X, Z, \alpha) p(\alpha|x, X, Z) \mathrm{d}w \mathrm{d}\alpha, \tag{3.40}$$

where the distribution is marginalized with respect to the weight vector $w \in \mathbb{R}^M$ and the hyperparameters $\alpha \in \mathbb{R}^M$. Since $\sigma(\cdot)$ is nonlinear, no closed-form solution exists for (3.40) [139]. Here, we use the *type 2 maximum likelihood* [8]—also known as the *evidence approximation* [101]—to approximate (3.40) by replacing $\alpha \in \mathbb{R}^M$ with a constant value $\alpha^* \in \mathbb{R}^M$ corresponding to the *mode* (i.e., the maximizer) of the *marginal likelihood function* $p(Z|X, \alpha)$. In particular, an approximation to the predictive distribution $p(z|x, X, Z)$ is given by

$$p(z|x, X, Z) \simeq p(z|x, X, Z, \alpha^*) = \int_{-\infty}^{\infty} p(z|x, X, Z, w, \alpha^*) p(w|x, X, Z, \alpha^*) \mathrm{d}w. \tag{3.41}$$

The value of α^* is found via an iterative process. After initializing α, the posterior distribution $p(w|x, X, Z, \alpha)$ is approximated by a Gaussian distribution using the Laplace approximation method. The mean of the Gaussian distribution corresponds to the mode (maximizer) of $p(w|x, X, Z, \alpha)$, which we denote by w^*. The maximizer is found using the iterative reweighted least squares (IRLS) method [15], which uses

sequential quadratic approximations to find the maximizer. Taking the log of the identity [15]

$$p(w|x, X, Z, \alpha) = \frac{p(Z|x, X, w, \alpha)p(w|x, X, \alpha)}{p(Z|x, X, \alpha)} = \frac{p(Z|X, w)p(w|\alpha)}{p(Z|X, \alpha)},$$

the maximization problem is equivalent to

$$\max_{w\in\mathbb{R}^M} \left\{\ln\left(p(Z|X, w)p(w|\alpha)\right) - \ln p(Z|X, \alpha)\right\}, \tag{3.42}$$

or, equivalently,

$$\max_{w\in\mathbb{R}^M} \left\{\sum_{n=1}^{N} \left(z_n \ln y_n + (1 - z_n)\ln(1 - y_n)\right) - \frac{1}{2}w^{\mathrm{T}}Aw + c\right\}, \tag{3.43}$$

where $y_n = \sigma\left(y(x_n)\right) \in \mathbb{R}$, $A = \mathrm{diag}[\alpha] \in \mathbb{R}^{M\times M}$, $c \in \mathbb{R}$ is a variable independent of z (and hence, plays no role in the optimization), and where we have used (3.38). Note that the covariance matrix of the Gaussian approximation to the posterior distribution $p(w|x, X, Z, \alpha)$ is equal to the negative Hessian of $\ln p(w|x, X, Z, \alpha)$ evaluated at the maximizer w^*. The mean and covariance of the Gaussian approximation are given by

$$w^* = A^{-1}\Phi^{\mathrm{T}}(Z - Y), \tag{3.44}$$

$$\Sigma = \left(\Phi^{\mathrm{T}}B\Phi + A\right)^{-1}, \tag{3.45}$$

where $\Phi = [\Phi_{(i,j)}] \in \mathbb{R}^{N\times M}$ with $\Phi_{(i,j)} = \phi_j(x_i)$ for $i = 1, \ldots, N$ and $j = 1, \ldots, M$, $B = \mathrm{diag}[b_1, b_2, \ldots, b_N] \in \mathbb{R}^{N\times N}$ with $b_n = y_n(1 - y_n) \in \mathbb{R}$, $n = 1, \ldots, N$, and $Y = [y_1, y_2, \ldots, y_N]^{\mathrm{T}} \in \mathbb{R}^N$.

Next, using (3.36), we approximate the marginal likelihood function as

$$\begin{aligned} p(Z|X, \alpha) &= \int_{-\infty}^{\infty} p(Z|X, w, \alpha)p(w|X, \alpha)\mathrm{d}w \\ &= \int_{-\infty}^{\infty} p(Z|X, w)p(w|\alpha)\mathrm{d}w \\ &\simeq p(Z|X, w^*)p(w^*|\alpha)(2\pi)^{N/2}(\det\Sigma)^{\frac{1}{2}}. \end{aligned} \tag{3.46}$$

Table 3.1: Sparse Bayesian Learning Algorithm

Step 1.	Initial parameter computation.
a.	Set $\Phi_{(i,j)} = \phi_j(x_i)$, $i = 1, \ldots, N$, $j = 1, \ldots, M$.
b.	Set $\Phi = [\Phi_{(i,j)}] \in \mathbb{R}^{N \times M}$.
c.	Set $b_n = y_n(1 - y_n)$, $n = 1, \ldots, N$.
d.	Set $B = \mathrm{diag}[b_1, b_2, \ldots, b_N]$.
Step 2.	Initialize the hyperparameters $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_M]^{\mathrm{T}}$.
Step 3.	Find the Gaussian approximation $\mathcal{N}(w; w^*, \Sigma)$ to the posterior distribution $p(w\|x, X, Z, \alpha)$.
a.	Set $A = \mathrm{diag}[\alpha]$.
b.	Set $w^* = A^{-1}\Phi^{\mathrm{T}}(Z - Y)$.
c.	Set $\Sigma = \left(\Phi^{\mathrm{T}} B \Phi + A\right)^{-1}$.
Step 4.	Compute the approximate marginal likelihood function using $p(Z\|X, \alpha) \simeq p(Z\|X, w^*)p(w^*\|\alpha)(2\pi)^{N/2}(\det \Sigma)^{\frac{1}{2}}$.
Step 5.	Set $\gamma_n = 1 - \alpha_n \Sigma_{nn}$, $n = 1, \ldots, M$.
Step 6.	Update α using $\alpha_n \leftarrow \frac{\gamma_n}{(w_n^*)^2}$, $n = 1, \ldots, M$.
Step 7.	If $\|\Delta\alpha\| > \mathrm{Tol}_1$ or $\|\Delta w^*\| > \mathrm{Tol}_2$, where $\Delta\alpha$ and Δw^* are the changes in the values of α and w^* in the current iteration, respectively, and Tol_1 and Tol_2 are some pre-specified tolerances, then go to Step 3.
Step 8.	Set $\alpha^* = \alpha$.

Following the discussion on the type 2 maximum likelihood method, the value of α^* is found by maximizing the approximate marginal likelihood function given by (3.46). Hence, differentiating (3.46) with respect to α_n, $n = 1, \ldots, M$, and setting the result to zero yields

$$-\frac{1}{2}(w_n^*)^2 + \frac{1}{2\alpha_n} - \frac{1}{2}\Sigma_{(n,n)} = 0, \quad n = 1, \ldots, M. \tag{3.47}$$

Solving (3.47) for α_n, $n = 1, \ldots, M$, gives the updated estimate for α_n as

$$\alpha_n^{\mathrm{new}} = \frac{\gamma_n}{(w_n^*)^2}, \quad n = 1, \ldots, M, \tag{3.48}$$

where $\gamma_n = 1 - \alpha_n \Sigma_{(n,n)}$, $n = 1, \ldots, M$. Now, using the updated estimate $\alpha^{\mathrm{new}} = [\alpha_1^{\mathrm{new}}, \alpha_2^{\mathrm{new}}, \ldots, \alpha_M^{\mathrm{new}}]^{\mathrm{T}} \in \mathbb{R}^M$ for α, the above steps are repeated until a given convergence criterion is met. The algorithm is summarized in Table 3.1.

As a result of the maximization of the marginal likelihood function, a number of

hyperparameters α_n approach infinity, and hence, the corresponding weight parameter w_n will be centered at zero with zero variance. Therefore, the corresponding component of the feature space transformation $\phi_n(\cdot)$ plays no role in the prediction, resulting in a sparse predictive model. In the case of the RVM, the input variables x_n belonging to the training set $\{x_1, x_2, \ldots, x_N\}$, which have a non-zero weight w_n, are called *relevance vectors*. Only the relevance vectors play a role in the classification of new input variables and the rest of the training set can be discarded.

Finally, we note that the posterior probability for the membership of a new input variable x to the class $\mathcal{C}_1$ represented by $p(\mathcal{C}_1|x, X, Z)$ can be approximated by the logistic sigmoidal function $\sigma(y(x))$ using the calculated value of the weight vector w^*. This approximatation becomes exact as the number of input variables in the training set approaches infinity [14, 139].

3.4. Pain and Pain Intensity Assessment in Neonates

In this section, we use the classification techniques described in Sections 3.2 and 3.3 in order to assess pain and pain intensity in infants using their facial expressions. For our data set we use the Infant Classification of Pain Expressions (COPE) database [22]. As was shown in [22], the SVM can classify facial images into two groups of "pain" and "non-pain" with an accuracy between 82% to 88%. Here we extend the results of [22] to additionally assess pain intensity using the class membership posterior probability. Note that although we consider infants, studies have shown that the pain-induced facial expressions in newborns are similar to those observed in older children and adults [37]. However, neonatal facial expressions are characterized by some unique features that are not found in adults such as "primal face of pain" [126]. In addition, adults can control nonverbal expressions of pain [64].

Figure 3.1: Four different expressions of a subject. The 2 left images correspond to non-pain, whereas the 2 right images correspond to pain.

Before applying the classification techniques to the facial images, we give a brief description of the infant COPE database used in our experimental results.

3.4.1. Infant COPE Database

The infant COPE database is composed of 204 RGB color photographs of 26 Caucasian neonates (13 boys and 13 girls) with a resolution of 120×100 per photograph and an infant age range of 18 hours to 3 days. The photographs were taken after a series of stress-inducing stimuli were administered by a nurse. The stimuli consist of the following [22]:

i) Transport from one crib to another.

ii) Air stimulus, where the infant's nose was exposed to a puff of air.

iii) Friction, where the external lateral surface of the heel was rubbed with a cotton wool soaked in alcohol.

iv) Pain, where the external surface of the heel was punctured for blood collection.

The facial expressions induced by the first three stimuli are classified as non-pain. Four photographs of a typical subject are given in Figure 3.1. One of the challenges in the recognition of pain, even for clinicians, is the ability to distinguish an infant's cry induced by pain and some other non-painful stimulus.

3.4.2. Pain Recognition using Sparse Kernel Machine Algorithms

The classification techniques discussed in Section 3.2 were used to identify the facial expressions corresponding to pain. A total of 21 subjects from the infant COPE database were selected such that for each subject at least one photograph corresponded to pain and one to non-pain. The total number of photographs available for each subject ranged between 5 to 12, with a total of 181 photographs considered. We applied the leave-one-out method for validation [15]. In particular, the classifier is trained on all photographs of the subject except for one test photograph which is used to validate the algorithm. The test photograph corresponds to either a pain or non-pain condition.

In the preprocessing stage, the faces were standardized for their eye position using a similarity transformation. Then, a 70×93 window was used to crop the facial region of the image and only the 8-bit grayscale values were used. For each image, a 6510-dimensional vector was formed by column stacking the matrix of intensity values.

We used the MATLAB® version R2008a and the OSU SVM MATLAB® Toolbox [116] to run the SVM classification algorithm. The classification accuracy for the SVM algorithm with a linear kernel was 90%, where, as suggested in [22], we chose the complexity parameter $C = 1$. The number of support vectors averaged 5. Applying the RVM algorithm with a linear kernel to the same data set resulted in an almost identical classification accuracy, namely, 91%; whereas the number of relevance vectors was reduced to 2. However, in 5 out of the 21 subjects considered, the RVM algorithm did not converge. This is due to the fact that, in contrast to the SVM algorithm, the RVM algorithm involves a non-convex optimization problem [15].

3.4.3. Pain Intensity Assessment

In addition to classification, the RVM algorithm provides the posterior probability of the membership of a test image to a class. As discussed in the introduction, using a Bayesian interpretation of probability, the probability of an event can be interpreted as the degree of the uncertainty associated with such an event. This uncertainty can be used to estimate pain intensity. In particular, if a classifier is trained with a series of facial images corresponding to pain and non-pain, then there is some uncertainty for associating the facial image of a person experiencing moderate pain to the pain class. The efficacy of such an interpretation of the posterior probability was validated by comparing the algorithm's pain assessment with that assessed by several experts (intensivists) and non-experts.

In order to compare the pain intensity assessment given by the RVM algorithm with human assessment, we compared the subjective measurement of the pain intensity assessed by expert and non-expert examiners with the uncertainty in the pain class membership (posterior probability) given by the RVM algorithm. Actual pain modes for each infant were used to train the RVM classifier. We chose all 16 infants (out of the total 21) from the COPE database for which the RVM algorithm converged, and for each subject two photographs of the face corresponding to the non-pain and pain conditions were selected. In the selection process, two training photographs were selected where the infant's facial expression truly reflected the pain intensity condition—calm for non-pain and distressed for pain—and a score of 0 and 100, respectively, was assigned to these photographs to give the human examiner a fair prior knowledge for the assessment of the pain intensity.

Ten data examiners were asked to provide a score ranging from 0 to 100 for each test photograph (i.e., non-training photograph) of the same subject, using a multiple

of 10 for the scores. Five examiners with no medical expertise and five examiners with medical expertise in critical care and pain management were selected for this assessment. The medical experts were members of the clinical staff at the intensive care unit of the Northeast Georgia Medical Center, Gainesville, GA, consisting of one medical doctor, one nurse practitioner, and three nurses. The medical doctor has 10 years experience as an anesthesiologist and intensivist in pediatric hospitals. The nurse practitioner and nurses have 6 months to 1 year pediatric floor experience in large community hospitals and are also mothers. They were asked to assess the pain for a series of random photographs of the same subject, with the criterion that a score above 50 corresponds to pain, and with an increasing score corresponding to a higher pain intensity. Analogously, a score below 50 corresponds to non-pain, and with a decreasing score corresponding to a lower level of discomfort. The posterior probability given by the RVM algorithm with a linear kernel for each corresponding photograph was rounded off to the nearest multiple of 10.

The pain scores for 5 infant subjects are given in Figures 3.2 – 3.6, where the average score of the expert and non-expert human examiners are compared to the score given by the RVM algorithm. In order to measure the agreement between the human examiners and the RVM algorithm, we need to quantify the agreement between two raters classifying an observation into different classes. The *kappa coefficient* [35] is used to measure the agreement between two raters classifying the same observation into two classes. A kappa coefficient of 0 represents chance agreement and a coefficient of 1 represents a perfect agreement between the two raters. The *weighted kappa coefficient* is an extension of the kappa coefficient to the case where there are more than two classes and the classes are ordered [36]. In this case, a smaller difference between the chosen classes by the two raters indicates less disagreement between them. The pain intensity assessment can be regarded as a classification process in

Table 3.2: Qualitative evaluation of the observed κ-values [96]

κ	Strength of agreement
0.00	Poor
0.00–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

which a facial expression of a subject is classified into 10 ordered classes, where Class 1 corresponds to a pain intensity score of $0-9$, Class 2 corresponds to a pain intensity score of $10-19$, and so on. A qualitative evaluation of the observed kappa values is given in Table 3.2 [46].

We used the weighted kappa coefficient to measure the agreement in the pain intensity assessment between the human examiners and the RVM algorithm. This coefficient is 0.47 for human expert examiners (with a 95% confidence interval of 0.37 to 0.57) and 0.46 for non-expert examiners (with a 95% confidence interval of 0.36 to 0.55) as compared with the RVM for the 16 subjects considered in the study. This shows a moderate agreement between the human expert examiners and human non-expert examiners as compared with the RVM algorithm based on the qualitative evaluation of the observed kappa values given by Table 3.2. It is interesting to note that the weighted kappa coefficient measuring the agreement between human experts and human non-experts is 0.78 with a 95% confidence interval of 0.73 to 0.82, which indicates a substantial agreement based on Table 3.2. It is important to note, however, that proxy ratings of pain is a highly subjective process [59].

The results show an almost identical classification accuracy for a binary classification (with a score above 50 corresponding to pain). In particular, the non-expert human examiner, the expert human examiner, and the RVM classification accuracy is

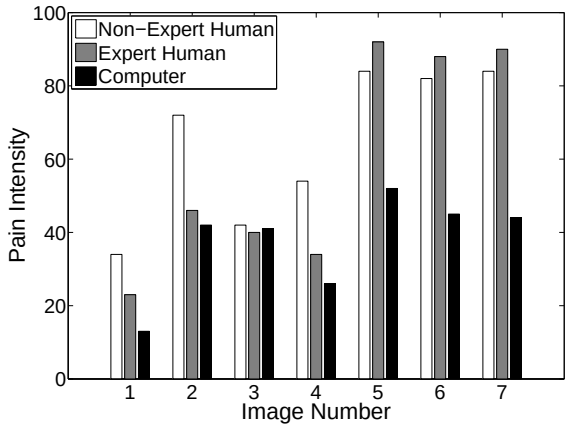

Figure 3.2: Pain score for Subject 1

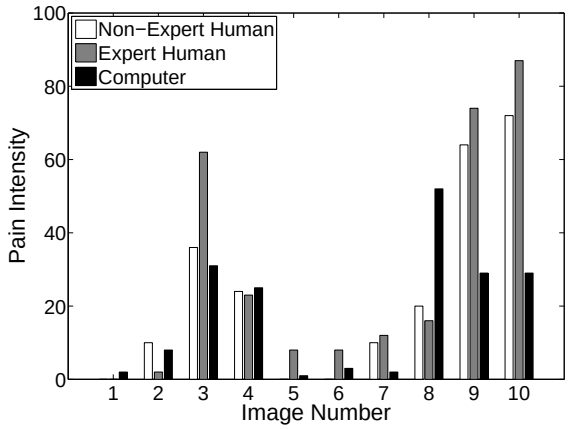

Figure 3.3: Pain score for Subject 2

given by 79%, 87%, and 91%, respectively. Moreover, the results show that the expert human examiners tend to be more accurate in the binary classification compared to the human non-experts.

It is worth noting that Figure 3.4 shows a poor correlation between the scores given by the RVM algorithm and the data examiners in the first three photographs. The data examiners assessed a high level of pain for Subject 3, whereas the subject was not in pain. This highlights the challenge in distinguishing between pain from discomfort, even for human experts. In this case, the RVM algorithm correctly assessed that the

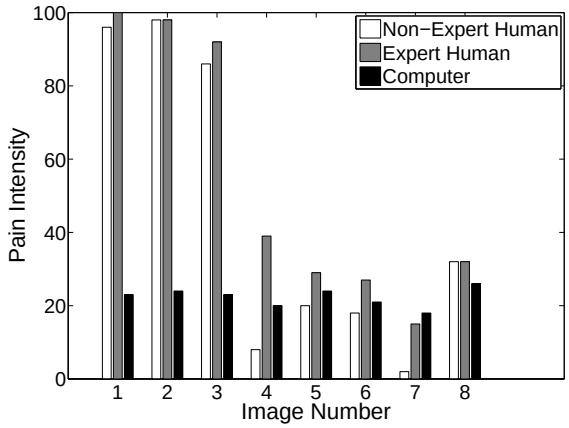

Figure 3.4: Pain score for Subject 3

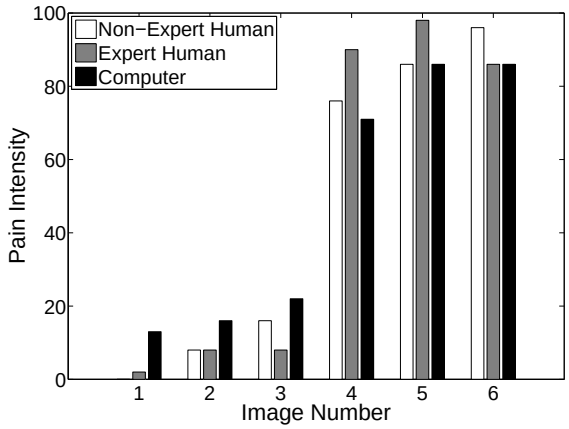

Figure 3.5: Pain score for Subject 4

infant has some level of discomfort but is not in pain.

Finally, in a repeatability study, the same human expert and non-expert examiners were asked to assess the intensity of pain for the 5 subjects considered in Figures 3.2 – 3.6 after a period ranging from 2 weeks to 4 months. Again, we used the weighted kappa coefficient to measure the agreement between two observations by the same rater. The weighted kappa coefficient in this case can be regarded as a measure of the ability of the human examiner to reproduce his or her own pain scores. The weighted kappa coefficient is 0.79 (with a 95% confidence interval of 0.74 to 0.84) for the human

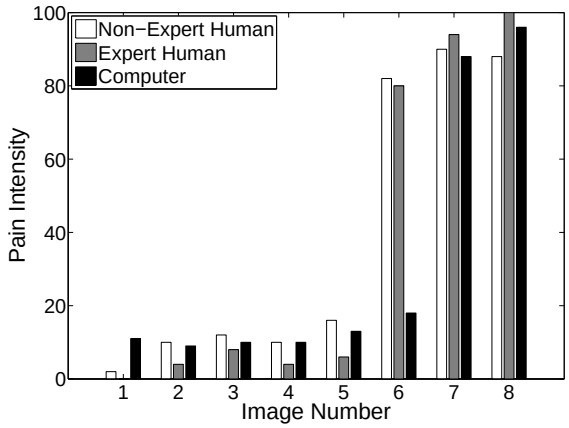

Figure 3.6: Pain score for Subject 5

expert examiners and 0.73 (with a 95% confidence interval of 0.68 to 0.78) for the human non-expert examiners. Based on this analysis, the human expert examiners tend to be slightly more reliable in assessing the pain intensity for the same subjects under the same pain conditions.

Chapter 4

An Unsupervised Learning Approach for Facial Expression Recognition using Semi-Definite Programming and Generalized Principal Component Analysis

4.1. Introduction

The human face is a rich medium through which people communicate their emotions. Researchers have identified six basic human expressions, namely, happiness, sadness, anger, disgust, fear, and surprise [45]. Automatic facial expression recognition algorithms can be used in systems involving human-computer interaction [117]. An emerging field of application for facial expression recognition algorithms involves clinical decision support systems [22, 60]. Specifically, the authors in [53, 54] present a framework for assessing pain and pain intensity in neonates using digital imaging.

Among different approaches proposed for facial expression recognition are *manifold-based methods* [30]. In these methods, the facial image can be regarded as a point in a D-dimensional space (which is referred to as the *ambient space*), where D is the number of pixels in the image or the number of parameters in a face model. The underlying assumption in manifold-based methods is that a set of facial images of a subject, which are represented by a set of points in a high-dimensional ambient space, resides on an intrinsically low-dimensional manifold. Hence, an important part of the

facial expression recognition algorithm in such methods involve finding the manifold of facial expressions.

In this chapter, we propose an unsupervised learning approach to facial expression recognition, where we show that different facial expressions reside on distinct subspaces if the manifold of facial images is unfolded [148]. Specifically, we introduce a new manifold-based method, where we use a maximum variance unfolding (MVU) approach [148] to identify the low-dimensional manifold of facial images and unfold it. Next, generalized principal component analysis is used to fit a series of subspaces to the data points and associate each data point to a subspace. Data points that belong to the same subspace are shown to belong to the same facial expression.

The contents of the chapter are as follows. First, we review the MVU dimension reduction technique, which involves semi-definite programming and convex optimization. In Section 4.3, we review the generalized principal component analysis (GPCA). This framework uses algebro-geometric concepts to address the problem of data segmentation and subspace identification for a given set of data points. Finally, in Section 4.4, the MVU and GPCA methods are used to recognize facial expressions from a given set of images within an unsupervised learning framework.

The notation used in this chapter is fairly standard. Specifically, $\mathbb{Z}_+$ denotes the set of positive integers, $\mathbb{R}$ denotes the set of real numbers, $(\cdot)^{\mathrm{T}}$ denotes transpose, and $(\cdot)^\dagger$ denotes the Moore-Penrose generalized inverse. Furthermore, we write $\mathrm{tr}(\cdot)$ for the trace operator, $\mathcal{N}(\cdot)$ for null space, $\|\cdot\|$ for the Euclidean norm, and $\dim(\mathcal{S})$ for the dimension of a set $\mathcal{S} \subset \mathbb{R}^D$, where $D \in \mathbb{Z}_+$.

4.2. Manifold Unfolding and Dimension Reduction

In this section, we introduce the method of maximum variance unfolding (MVU), which involves a dimension reduction technique that uses semi-definite programming and convex optimization. Given a set of points sampled from a low-dimensional manifold in a high-dimensional ambient space, this technique unfolds the manifold (and hence, the points it contains) while preserving the local geometrical properties of the manifold [148]. This method can be regarded as a nonlinear generalization of the principal component analysis (PCA) [148].

Given a set of points in a high-dimensional ambient space, principal component analysis identifies a low-dimensional subspace such that the variance of the projection of the points on this subspace is maximized. More specifically, the basis of the subspace on which the projection of the points has the maximum variance is the eigenvectors corresponding to the non-zero eigenvalues of the covariance matrix [88]. In the case where the data is noisy, the singular vectors corresponding to the dominant singular values of the covariance matrix are selected [145, 146].

Given the set of N input points $\mathcal{X} = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$, where D is the dimension of the ambient space, we seek to find the set of N output points $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^d$ such that $d < D$, $\mathcal{X}$ and $\mathcal{Y}$ are equivalent, and points sufficiently close to each other in the input data set $\mathcal{X}$ remain sufficiently close in the output data set $\mathcal{Y}$. Recall that two sets $\mathcal{X}$ and $\mathcal{Y}$ are *equivalent* if and only if there exists a bijective (one-to-one and onto) map $f : \mathcal{X} \to \mathcal{Y}$. To address this problem, the concept of *isometry* for a set of points is needed [136, 148]. In particular, an isometry is a diffeomorphism defined on a manifold such that it admits a local translation and rotation. The next definition extends the notion of isometry to data sets.

Definition 4.1 [148]. Let $\mathcal{X} = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$ and $\mathcal{Y} = \{y_1, y_2, \ldots,$

$y_N\} \subset \mathbb{R}^d$ be equivalent. Then $\mathcal{X}$ and $\mathcal{Y}$ are *k-locally isometric* if there exists a continuous map $T : \mathbb{R}^D \to \mathbb{R}^d$ such that if $T(x_i) = y_i$, then $T(N_{x_i}(k)) = N_{y_i}(k)$, $i = 1, \ldots, N$, where $N_{x_i}(k)$ (resp., $N_{y_i}(k)$) is the set of k-nearest neighbors of $x_i \in \mathcal{X}$ (resp., $y_i \in \mathcal{Y}$).

Before stating the MVU method, we introduce the following maximization problem.

Maximum Variance Unfolding Problem. Given a set of input data points $\mathcal{X} = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$ find the set of output data points $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^d$, where $d \leq D$, such that the sum of pairwise square distances between the outputs in $\mathcal{Y}$ given by

$$\Phi = \frac{1}{2N} \sum_{i=1}^{N} \sum_{j=1}^{N} \|y_i - y_j\|^2, \tag{4.1}$$

is maximized, and $\mathcal{X}$ and $\mathcal{Y}$ are k-locally isometric for some $k \in \mathbb{Z}_+$.

Without loss of generality, we assume that $\sum_{i=1}^{N} x_i = 0$. Moreover, we require $\sum_{n=1}^{N} y_n = 0$ to remove the translational degree of freedom in the output points contained in $\mathcal{Y}$. Note that a data set (e.g., $\mathcal{X}$) can be represented by a weighted graph $\mathfrak{G}$ [138], where each node represents a point and the k-nearest points are connected by edges, where k is a given parameter. The weights of $\mathfrak{G}$ represent the distance between the nodes. In addition, we assume that the graph $\mathfrak{G}$ is connected [138]. In the case of a disconnected graph, each connected component should be analyzed separately. The k-local isometry condition in the Maximum Variance Unfolding Problem requires that the distances and the angles between the k-nearest neighbors are preserved. This constraint is equivalent to preserving the distances between neighboring points in a modified graph $\mathfrak{G}'$, where, for each node, all the neighboring nodes of $\mathfrak{G}'$ are connected by an edge. More precisely, each node of $\mathfrak{G}'$ and the k-neighboring nodes of $\mathfrak{G}'$ form a *clique* of size $k + 1$ (see Figure 4.1) [138].

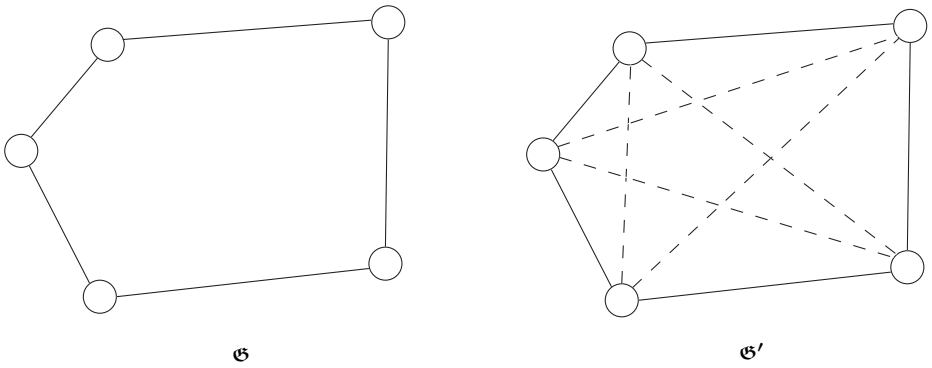

Figure 4.1: Original and modified graphs for $k = 2$.

The next theorem gives the solution to Maximum Variance Unfolding Problem for the case where $d = D$.

Theorem 4.1 [148]. Consider the Maximum Variance Unfolding Problem with $d = D$. The output data points in $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$ are given by the solution to the optimization problem

$$\max_{y_1, y_2, \ldots, y_N \in \mathbb{R}^D} \Phi, \tag{4.2}$$

subject to

$$\sum_{i=1}^{N} y_i = 0, \tag{4.3}$$

$$\|y_i - y_j\|^2 = D_{ij}, \quad \text{if } \eta_{(i,j)} = 1, \quad i, j = 1, \ldots, N, \tag{4.4}$$

where Φ is defined in (4.1), $\eta = [\eta_{(i,j)}] \in \mathbb{R}^{N \times N}$ is the adjacency matrix of the modified graph $\mathfrak{G}'$, and

$$D_{ij} = \|x_i - x_j\|^2, \quad i, j = 1, \ldots, N, \quad x_i, x_j \in \mathcal{X}. \tag{4.5}$$

The optimization problem (4.2)–(4.4) is *not* convex. The following convex optimization problem, however, is equivalent to the optimization problem given in Theorem 4.1. Moreover, the following result also addresses the case where $d \leq D$.

Theorem 4.2 [148]. Consider the Maximum Variance Unfolding Problem with $d = D$ and let $\mathfrak{G}$ and $\mathfrak{G}'$ denote the weighted graph and modified graph corresponding to the data set $\mathcal{X}$, respectively. The output data points in $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$ are given by

$$y_{ji} = \sqrt{\lambda_j V_{ji}}, \quad j = 1, \ldots, N, \quad i = 1, \ldots, D, \tag{4.6}$$

where V_{ji}, $j = 1, \ldots, N$, $i = 1, \ldots, D$, is the ith component of the jth eigenvector of K^* given by $V_j = [V_{j1}, V_{j2}, \ldots, V_{jD}]^{\mathrm{T}}$, λ_j is the associated eigenvalue, y_{ji}, $j = 1, \ldots, N$, $i = 1, \ldots, D$, is the ith component of the vector $y_j = [y_{j1}, y_{j2}, \ldots, y_{jD}]^{\mathrm{T}}$, and K^* is the optimal solution to the optimization problem

$$\max_{K \in \mathbb{K}} \operatorname{tr}(K), \tag{4.7}$$

subject to

$$K \geq 0, \tag{4.8}$$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} K_{(i,j)} = 0, \tag{4.9}$$

$$K_{(i,i)} - 2K_{(i,j)} + K_{(j,j)} = D_{ij}, \quad \text{if } \eta_{(i,j)} = 1, \quad i, j = 1, \ldots, N, \tag{4.10}$$

where $\mathbb{K} \subset \mathbb{R}^{N \times N}$ denotes the cone of nonnegative definite matrices, $\eta = [\eta_{(i,j)}] \in \mathbb{R}^{N \times N}$, and D_{ij} is defined as in (4.5). Moreover, if K^* has $d < D$ non-zero eigenvalues, then the set of reduced dimension output data points is given by $\mathcal{Y} = \{y_1^{\mathrm{red}}, y_2^{\mathrm{red}}, \ldots, y_N^{\mathrm{red}}\} \subset \mathbb{R}^d$, where y_i^{red}, $i = 1, \ldots, N$, is found by removing the zero elements of y_i.

Remark 4.1. When the data is noisy, all the eigenvalues of K are typically non-zero, and hence, one has to choose the dominant eigenvalues of K^* [145, 146]. If the eigenvalues of K are sorted in descending order, then the first d components of y_i, $i = 1, \ldots, N$, form a d-dimensional data set that is *approximately* k-locally isometric to $\mathcal{X} = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{R}^D$ [148].

4.3. Data Segmentation and Subspace Identification

In this section, we address the problem of data segmentation and subspace identification for a given set of data points. First, we define the multiple subspace segmentation problem.

Data Segmentation and Subspace Identification Problem [145,146]. Given the set $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$, where y_i, $i = 1, \ldots, N$, are drawn from a set of distinct subspaces $\mathcal{S}_j$, $j = 1, \ldots, n$, of unknown number and dimension, find *i*) the number of subspaces n, *ii*) their dimensions $\dim(\mathcal{S}_j)$, and *iii*) the basis for each subspace. Furthermore, associate each point to the subspace it belongs to.

GPCA uses algebro-geometric concepts to address this problem. First, we present the GPCA algorithm followed by a more robust version of GPCA to deal with noisy data. For a detailed discussion, see [145, 146].

4.3.1. Generalized Principal Component Analysis

In this subsection, we present the GPCA algorithm, where we assume that the data points are noise-free. The GPCA algorithm consists of three main steps; namely, polynomial fitting, polynomial differentiation, and polynomial division. The following definitions are needed.

Definition 4.2 [44, 51, 135, 145]. Let $\mathcal{R}$ be a commutative ring and let $\mathcal{I}$ be an additive subgroup of $\mathcal{R}$. $\mathcal{I}$ is called an *ideal* if $r \in \mathcal{R}$ and $s \in \mathcal{I}$, then $rs \in \mathcal{I}$. Furthermore, an ideal is said to be *generated* by a set $\mathcal{S}$ if, for all $t \in \mathcal{I}$, $t = \sum_{i=1}^{n} r_i s_i$, $r_i \in \mathcal{R}$, $s_i \in \mathcal{S}$, $i = 1, \ldots, n$, for some $n \in \mathbb{Z}_+$. Let $\mathbb{F}[x]$ be the set of polynomials of D variables, where $x = [x_1, x_2, \ldots, x_D]^{\mathrm{T}}$, $x_i \in \mathbb{F}$, $i = 1, \ldots, D$, and $\mathbb{F}$ is a field. Then $\mathbb{F}[x]$, with polynomial addition and multiplication, forms a commutative ring.

A product of n variables $x_1, x_2, \ldots, x_n$ is called a *monomial of degree* n (counting multiplicity). The number of distinct monomials of degree n is given by

$$M_n(D) \triangleq C(D+n-1, n), \tag{4.11}$$

where $C(p,q)$ denotes the combination of p objects taken q at a time. A polynomial with all of its terms being the same degree is called a *homogenous polynomial*. An ideal that is generated by homogenous polynomials is called a *homogenous ideal*. Finally, the *Veronese Map of degree* n is a mapping $\nu_n : \mathbb{F}^D \rightarrow \mathbb{F}^{M_n(D)}$ that assigns to the variables $x_1, x_2, \ldots, x_D$ all the possible monomials of degree n; namely,

$$\nu_n([x_1,\, x_2,\, \ldots,\, x_D]^{\mathrm{T}}) = [u_1,\, u_2,\, \ldots,\, u_{M_n(D)}]^{\mathrm{T}}, \tag{4.12}$$

where $u_i = x_1^{n_{i1}} x_2^{n_{i2}} \cdots x_D^{n_{iD}}$, $i = 1, \ldots, M_n(D)$, and where $n_{i1} + n_{i2} + \ldots + n_{iD} = n$, $n_{ij} \in \mathbb{Z}_+$, $j = 1, \ldots, D$, and $n_{i1}, \ldots, n_{iD}$ are in lexicographic order.

Let $\mathcal{A} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n\}$, $\mathcal{Z}_{\mathcal{A}} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \ldots \cup \mathcal{S}_n$, where $\mathcal{S}_j \subset \mathbb{R}^D$, $j = 1, \ldots, n$, is a linear subspace, $\dim(\mathcal{S}_j) = d_j$, and $n \in \mathbb{Z}_+$. $\mathcal{A}$ is referred to as a *subspace arrangement*. In addition, let $\mathcal{Y} = \{y_1,\, y_2, \ldots,\, y_N\} \subset \mathbb{R}^D$ be a set of a sufficiently large number of points sampled from $\mathcal{Z}_{\mathcal{A}}$. In this chapter, we assume that the number of subspaces n is known. However, the GPCA algorithm, in its most general form, gives the solution for the case where n is unknown [145, 146]. In order to algebraically represent $\mathcal{Z}_{\mathcal{A}}$, we need to find the *vanishing ideal* of $\mathcal{Z}_{\mathcal{A}}$, denoted by $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$. The vanishing ideal of $\mathcal{Z}_{\mathcal{A}}$ is the set of polynomials which vanish on $\mathcal{Z}_{\mathcal{A}}$. It can be shown that the homogenous component of $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$, denoted by $\mathcal{I}_n$, uniquely determines $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$. Hence, to find the vanishing ideal $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$ it suffices to determine the homogenous component $\mathcal{I}_n$.

If $p_n(x)$ is a polynomial in $\mathcal{I}_n$, then $p_n(x) = c_n^{\mathrm{T}} \nu_n(x)$, where $c_n \in \mathbb{R}^{M_n(D)}$, $\nu_n(x) : \mathbb{F}^D \rightarrow \mathbb{F}^{M_n(D)}$ is the veronese map given by (4.12), $x = [x_1,\, x_2,\, \ldots,\, x_D]^{\mathrm{T}}$ for some

$D \in \mathbb{Z}_+$, and $M_n(D)$ is given by (4.11). Therefore, every point y_i, $i = 1, \ldots, N$, satisfies $p_n(x) = 0$, and hence, $V_n(D)c_n = 0$, where

$$V_n(D) \triangleq \begin{bmatrix} \nu_n^{\mathrm{T}}(y_1) \\ \nu_n^{\mathrm{T}}(y_2) \\ \vdots \\ \nu_n^{\mathrm{T}}(y_N) \end{bmatrix} \tag{4.13}$$

is called the *embedded data matrix*. A one-to-one correspondence between the null space of $V_n(D)$ and the polynomials in $\mathcal{I}_n$ exists if

$$\dim(\mathcal{N}(V_n(D))) = \dim(\mathcal{I}_n) = h_{\mathcal{I}}(n), \tag{4.14}$$

or, equivalently,

$$\operatorname{rank} V_n(D) = M_n(D) - h_{\mathcal{I}}(n), \tag{4.15}$$

where the *Hilbert function* $h_{\mathcal{I}}(n)$ is the number of linearly independent polynomials of degree n that vanish on $\mathcal{Z}_{\mathcal{A}}$. The singular vectors of $V_n(D)$ denoted by $c_{ni} \in \mathbb{R}^{M_n(D)}$, $i = 1, \ldots, h_{\mathcal{I}}(n)$, corresponding to the zero singular values of $V_n(D)$ can be used to compute a basis for $\mathcal{I}_n$, namely

$$\mathcal{I}_n = \operatorname{span}\{p_{ni}(x) = c_{ni}\nu_n(x),\, i = 1, \ldots, h_{\mathcal{I}}(n)\}.$$

In the case where the data set $\mathcal{Y}$ is corrupted by noise, the singular vectors corresponding to the $h_{\mathcal{I}}(n)$ smallest singular values of $V_n(D)$ are used to compute the basis for $\mathcal{I}_n$.

The following theorem shows how polynomial differentiation can be used to find the dimensions and bases of each subspace $\mathcal{S}_j$, $j = 1, \ldots, n$.

Theorem 4.3 [145, 146]. Let $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$ be a set of points sampled from $\mathcal{Z}_{\mathcal{A}} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \ldots \cup \mathcal{S}_n$, where, for $j = 1, \ldots, n$, $\mathcal{S}_j$ is a subspace of unknown dimension d_j. Furthermore, assume that for every subspace $\mathcal{S}_j$, $j = 1, \ldots, n$, there

exists $w_j \in \mathcal{S}_j$ such that $w_j \notin \mathcal{S}_i$, $i \neq j$, $i = 1, \ldots, n$, and condition (4.14) holds. Then

$$\mathcal{S}_j^{\perp} = \text{span}\left\{ \frac{\partial}{\partial x} c_n^{\mathrm{T}} \nu_n(x)|_{x=w_j} : c_n \in \mathcal{N}\left(V_n(D)\right) \right\}, \quad j = 1, \ldots, n, \tag{4.16}$$

where $V_n(D)$ is the embedded data matrix of $\mathcal{Y}$ given by (4.13). Furthermore, $d_j = D - \text{rank}\, \nabla P_n(w_j)$, $j = 1, \ldots, n$, where $P_n(x) = [p_{n1}(x), p_{n2}(x), \ldots, p_{nh_{\mathcal{I}}(n)}(x)]^{\mathrm{T}} \in \mathbb{R}^{1 \times h_{\mathcal{I}}(n)}$ is a row vector of independent polynomials in $\mathcal{I}_n$, composed of the singular vectors corresponding to the zero singular values of $V_n(D)$, and $\nabla P_n = [\nabla^{\mathrm{T}} p_{n1}(x), \nabla^{\mathrm{T}} p_{n2}(x), \ldots, \nabla^{\mathrm{T}} p_{nh_{\mathcal{I}}(n)}(x)] \in \mathbb{R}^{D \times h_{\mathcal{I}}(n)}$.

To select a point w_j, $j = 1, \ldots, n$, for each subspace such that $w_j \in \mathcal{S}_j$, $w_j \notin \mathcal{S}_i$, $i \neq j$, $i = 1, \ldots, n$, without loss of generality, let $j = n$. One can show that the first point w_n, where $w_n \in \mathcal{S}_n$ and $w_n \notin \mathcal{S}_i$, $i = 1, \ldots, n-1$, is given by [145, 146]

$$w_n = \text{argmin}_{y \in \mathcal{Y}: \nabla P_n(y) \neq 0} P_n(y) (\nabla^{\mathrm{T}} P_n(y) \nabla P_n(y))^{\dagger} P_n^{\mathrm{T}}(y). \tag{4.17}$$

Furthermore, a basis for $\mathcal{S}_n$ can be found by applying PCA to $\nabla P_n(w_n)$. To find the rest of the points $w_i \in \mathcal{S}_i$, $i = 1, \ldots, n-1$, we can use polynomial division as outlined in the next theorem.

Theorem 4.4 [145, 146]. Let $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$ be a set of points sampled from $\mathcal{Z}_{\mathcal{A}} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \ldots \cup \mathcal{S}_n$, where, for $j = 1, \ldots, n$, $\mathcal{S}_j$ is a subspace of unknown dimension d_j. Assume (4.14) holds, $\mathcal{S}_n^{\perp}$ is known, and a point $w_n \in \mathcal{S}_n$ is given. Then, the set $\bigcup_{j=1}^{n-1} \mathcal{S}_j$ is characterized by the set of homogenous polynomials given by

$$\left\{ c_{n-1}^{\mathrm{T}} \nu_{n-1}(x) : V_n(D) R_n(b_n) c_{n-1} = 0,\, b_n \in \mathcal{S}_n^{\perp},\, c_{n-1} \in \mathbb{R}^{M_{n-1}(D)} \right\},$$

where $R_n(b_n) \in \mathbb{R}^{M_n(D) \times M_{n-1}(D)}$ is the matrix of coefficients of c_{n-1} when $(b_n^{\mathrm{T}} x)(c_{n-1}^{\mathrm{T}} \nu_{n-1}(x)) \equiv c_n^{\mathrm{T}} \nu_n(x)$ is rearranged to have the form $R_n(b_n) c_{n-1} = c_n$.

Once the homogenous polynomials $\{c_{n-1}^{\mathrm{T}}\nu_{n-1}(x)\}$ given by Theorem 4.4 are obtained, an identical procedure can be repeated to find w_{n-1} and the homogenous polynomials characterizing $\bigcup_{j=1}^{n-2}\mathcal{S}_j$.

4.3.2. Subspace Estimation Using a Voting Scheme

The GPCA framework given in Subsection 4.3.1 works well in the absence of noise. In practice, however, noise is always present and efficient statistical methods need to be incorporated with the GPCA. In this subsection, we present one such statistical method where a voting scheme is combined with the GPCA. Here, we assume that the number of the subspaces and their dimensions are known. For details, see [145, 146].

Let $\mathcal{Y} = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^D$ be the set of data points sampled from the set $\mathcal{Z}_{\mathcal{A}} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \ldots \cup \mathcal{S}_n$, where, for $j = 1, \ldots, n$, $\mathcal{S}_j$ is a subspace of dimension d_j and co-dimension $c_j = D - d_j$. As noted in Subsection 4.3.1, the homogenous component of degree n of the vanishing ideal $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$, denoted by $\mathcal{I}_n$, uniquely defines $\mathcal{I}(\mathcal{Z}_{\mathcal{A}})$ and $\dim(\mathcal{I}_n) = h_{\mathcal{I}}(n)$, where $h_{\mathcal{I}}(n)$ is the Hilbert function. Let $P = \{p_1(x), p_2(x), \ldots, p_{h_{\mathcal{I}}(n)}(x)\}$ be the set of polynomials forming a basis for $\mathcal{I}_n$. This set is obtained by selecting the $h_{\mathcal{I}}(n)$ smallest singular values of $V_n(D)$, where $V_n(D)$ is the embedded data matrix given by (4.13). Let $y_1 \in \mathcal{Y}$ and define $\nabla P_{\mathrm{B}}(y_1) \triangleq \left[\nabla^{\mathrm{T}} p_1(y_1), \nabla^{\mathrm{T}} p_2(y_1), \ldots, \nabla^{\mathrm{T}} p_{h_{\mathcal{I}}(n)}(y_1)\right]$. Note that in the noise-free case, if $y_1 \in \mathcal{S}_j$, then $\operatorname{rank} \nabla P_{\mathrm{B}}(y_1) = c_j$.

In the case where the data is corrupted by noise, a more efficient method for computing the basis is desired. Suppose the co-dimension of the subspaces $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ take q distinct values $c_1', c_2', \ldots, c_q'$, respectively. Given the fact that the membership of y_1 to one of the subspaces $\mathcal{S}_j$, $j = 1, \ldots, n$, is unknown, a set of basis candidates for the orthogonal complement of subspaces of all possible dimensions c_i', $i = 1, \ldots, q$, is calculated by choosing the c_i' principal components of $\nabla P_{\mathrm{B}}(y_1)$. This results in

q matrices $B_i \in \mathbb{R}^{D\times c_i'}$, $i = 1,\ldots,q$, each of which is a basis candidate for $\mathcal{S}_i^{\perp}$, $i = 1,\ldots,n$.

The main idea of the voting scheme is to count the number of repetitions of each basis candidate for all points in the data set $\mathcal{Y} = \{y_1,\ldots,y_N\}$. The n basis candidates with the most votes are chosen to be the basis for $\mathcal{S}_i^{\perp}$, $i = 1,\ldots,n$, and each point is assigned to its closest subspace. In our criterion for counting the repetition of the basis candidates, two basis candidates are considered to be the same if the angle between the subspaces spanned by them is less than τ, where $\tau > 0$ is a given tolerance parameter.

4.4. Unsupervised Learning of Facial Expressions

The MVU and GPCA methods presented in Sections 4.2 and 4.3 can be used to recognize facial expressions from a given set of images within an unsupervised learning framework. Specifically, given a set of images of a person with two different facial expressions (e.g., neutral and happy), we show that the two facial expressions reside on two distinct subspaces if the manifold is unfolded. In particular, semi-definite embedding is used to reduce the dimensionality and unfold the manifold of facial images. Next, generalized principal component analysis is used to fit a series of subspaces to the data points and associate each data point to a subspace. Data points that belong to the same subspace are claimed to belong to the same facial expression. The algorithm is summarized in Table 4.1.

In our experiment, 30 photographs were taken for each human subject, where the subject starts by a neutral expression, transitions to a happy expression, and goes back to a neutral expression with each part containing 10 photographs. An example set of images is given in Figure 4.2. These images were taken in a sequence, each

Table 4.1: Facial Expression Recognition Algorithm

Step 1.	Preprocess of the grayscale image data $I_1, \ldots, I_N$.
a.	Compute $x_j \in \mathbb{R}^{D'}$ by column stacking the matrix of I_j, $j = 1, \ldots, N$.
b.	Set the number of neighbors k.
c.	Form the weighted graph $\mathfrak{G}$.
d.	Form the modified graph $\mathfrak{G}'$ and the adjacency matrix $\eta = [\eta_{(i,j)}]$.
Step 2.	Manifold unfolding and dimension reduction.
a.	Set the dimension of the reduced space D.
b.	Find K^*, the maximizer of (4.7) subject to (4.8)–(4.10).
c.	Compute the eigenvectors and the associated eigenvalues of K^*; denote by $V_j = [V_{j1}, V_{j2}, \ldots, V_{jD'}]^{\mathrm{T}}$ and λ_j, $j = 1, \ldots, N$.
d.	Reorder V_j and λ_j such that λ_j, $j = 1, \ldots, N$ are in decreasing order.
e.	Compute $y_{ji} = \sqrt{\lambda_j V_{ji}}, \quad j = 1, \ldots, N, \quad i = 1, \ldots, D'$.
f.	Compute $y_j = [y_{j1}, y_{j2}, \ldots, y_{jD'}]^{\mathrm{T}}, \quad j = 1, \ldots, N$.
g.	Compute $y_j^{\mathrm{red}} = [y_{j1}, \ldots, y_{jD}]^{\mathrm{T}}, \quad j = 1, \ldots, N$.
h.	Compute $\mathcal{Y} = \{y_1^{\mathrm{red}}, \ldots, y_N^{\mathrm{red}}\}$.
Step 3.	Subspace estimation using a voting scheme.
a.	Set the subspace angle tolerance parameter $\tau > 0$.
b.	For the subspaces $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$, compute the distinct value of their co-dimension $c_1', c_2', \ldots, c_q'$.
c.	Initialize the arrays $\mathcal{U}_1 = [\,], \ldots, \mathcal{U}_q = [\,]$ and the counters $\mathcal{C}_1 = [\,], \ldots, \mathcal{C}_q = [\,]$
d.	**for** $j = 1 : N$
e.	**for** $i = 1 : q$
f.	Compute the c_i' principal components of $\nabla P_{\mathrm{B}}(y_j^{\mathrm{red}})$.
g.	Form the orthogonal matrix $B_i \in \mathbb{R}^{D \times c_i'}$ using outputs of **Step 3f.**
h.	**if** there exists k such that the angle $\angle_{\mathrm{subspace}}(B_i, \mathcal{U}_i(k)) < \tau$, **then**
i.	$\mathcal{C}_i(k) \leftarrow \mathcal{C}_i(k) + 1$.
j.	**else**
k.	$\mathcal{U}_i \leftarrow [\mathcal{U}_i, B_i]$.
l.	$\mathcal{C}_i \leftarrow [\mathcal{C}_i, 1]$.
m.	**end for**
n.	**end for**
o.	Select basis candidates from $\mathcal{U}_1, \ldots, \mathcal{U}_q$ corresponding to the n highest values of $\mathcal{C}_1, \ldots, \mathcal{C}_q$. Denote basis by $B_1, \ldots, B_n$.
p.	Use $B_1, \ldots, B_n$ (basis for $\mathcal{S}_1^{\perp}, \ldots, \mathcal{S}_n^{\perp}$) to find the $B_1', \ldots, B_n'$ (basis for $\mathcal{S}_1, \ldots, \mathcal{S}_n$).
q.	Use results in **Step 3p** to assign each $y_1^{\mathrm{red}}, \ldots, y_N^{\mathrm{red}}$, to the closest subspace.

200×240 pixels, and in total there were 4 subjects.

Each image can be considered as a vector of dimension 48000 by column stacking the grey scale image intensity matrix. In this case, each image is a point in a 48000-dimension space. In order to segment the images, the dimension is reduced to $D = 5$ using the MVU algorithm. Then, the resulting points in the $D = 5$-dimensional ambient space are used to identify 2 subspaces of dimension $d = 1, 2, 3, 4$, where in the GPCA voting algorithm two subspaces are considered to be the same if the angle between the two subspaces is less than $\tau = 0.4$ [152]. The segmentation error for each

case is given in Table 4.2, where it is noted that the best results correspond to $d = 3$ and 4. In order to visualize the subspace identification, the segmentation for the case $D = 2$ and $d = 1$ is given in Figure 4.3. Although these parameters result in a poor segmentation performance, it graphically conveys the main idea of the algorithm.

Figure 4.2: A sequence of pictures, where the subject starts with a neutral expression, smiles, and resumes to a neutral expression.

Table 4.2: Segmentation Results for $D = 5$

Subject	Number of Images	Segmentation Error			
		$d = 1$	$d = 2$	$d = 3$	$d = 4$
1	29	3	2	2	3
2	31	13	13	3	7
3	31	6	15	2	4
4	32	13	15	1	1

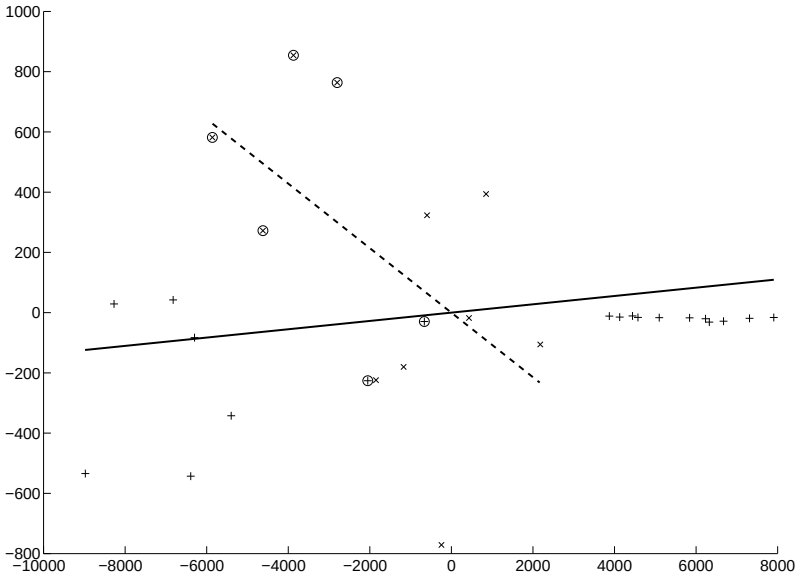

Figure 4.3: Facial expression segmentation with $D = 2$ and $d = 1$. The categorization error is 6/30. The solid and dashed lines are the subspaces corresponding to the neutral and happy expressions, respectively. The points associated with the solid line and the dashed line are depicted by "+" and "×", respectively. The points depicted by "○" are points associated with the wrong expression. Note that although these parameters result in a poor segmentation performance, it graphically conveys the main idea of the algorithm.

Chapter 5

Optimal Drug Dosing Control for Intensive Care Unit Sedation Using a Hybrid Deterministic-Stochastic Pharmacokinetic and Pharmacodynamic Model

5.1. Introduction

The clinical management of critically ill patients requiring mechanical ventilation due to respiratory failure is complex. Mechanical ventilation is intrinsically uncomfortable to the patient due to both the introduction of an artificial airway that is the interface between the patient and the ventilator, and also because of lack of synchronization between the patient's own spontaneous efforts to breathe and the action of the ventilator to breath for the patient. This can lead to the patient "fighting the ventilator" which is not only uncomfortable for the patient but can also have deleterious physiological effects. For this reason, patients often require administration of sedative and analgesic agents in ICUs.

In clinical ICU practice sedative/analgesic agents are titrated to achieve a specific level of sedation. The level of sedation is currently based on clinical scoring systems. Examples include the motor activity assessment scale (MAAS) [40], the Richmond agitation-sedation scale (RASS) [129], and the modified Ramsay sedation scale (MRSS) [121]. Specifically, in the MRSS scoring system patients are given an

integer score of 0–6 as follows: 0 - paralyzed, unable to evaluate; 1 - awake; 2 - lightly sedated; 3 - moderately sedated, follows simple commands; 4 - deeply sedated, responds to nonpainful stimuli; 5 - deeply sedated, responds only to painful stimuli; and 6 - deeply sedated, unresponsive to painful stimuli. In this chapter, we specifically consider the MRSS scoring system; however, the framework presented herein can be adopted to any other sedation scoring system. In addition, we assume that the patient's sedation level can always be evaluated, that is, the patient's MRSS sedation score of 1 to 6 can be assessed.

The goal of the clinician is to find the drug dose that maintains the patient at a sedation score of 3. This is typically done empirically, administering a drug dose that usually is in the effective range for most patients, observing the patient's response, and then adjusting the dose accordingly. However, the response of patients to any drug dose is a reflection of the pharmacokinetic and pharmacodynamic properties of the drug and the specific patient. In this chapter, we use pharmacokinetic and pharmacodynamic modeling to find an optimal drug dose, as a function of time, to drive the patient to an MRSS score of 3. This framework is developed for a general n-compartment mammillary phamacokinetic model and the methodology can be applied to any sedative agent.

Although pharmacokinetics of sedative and anesthetic drugs can be adequately modeled by nonnegative and compartmental dynamical systems [61], the pharmacodynamics of these drugs are not well understood and drug effect predictions usually involve probabilities [4, 5, 60]. Specifically, when considering sedative agents, drug effect is closely related to patient sedation level. As discussed in [4, 5], the corresponding sedation level of the ICU patient is related to drug concentration in the effect-site compartment using an empirical probabilistic model.

In this chapter, we model the pharmacokinetics and pharmacodynamics of a gen-

eral sedative agent using a hybrid deterministic-stochastic model involving deterministic pharmacokinetics and stochastic pharmacodynamics. Then, using this hybrid model, we consider the sedative drug propofol and use nonnegative and compartmental modeling to model the drug pharmacokinetics (drug concentration as a function of time) and a stochastic process to represent the patient's sedation score and model drug the pharmacodynamics (drug effect as a function of concentration). The first-order distribution of the stochastic process is a function of the states of the compartmental dynamical system.

Next, we use the aforementioned hybrid deterministic-stochastic model to develop an open-loop optimal control policy for ICU sedation. Specifically, we first find the optimal effect-site drug concentration corresponding to a high probability for the desired sedation score (i.e., MRSS score of 3) and a low probability for all other sedation scores. Then, we use optimal control theory to drive the effect-site drug concentration to the optimal value found in the previous step while minimizing a given cost functional. The cost functional captures control effort constraints as well as probability constraints associated with different sedation scores. The proposed methodology is then applied to a three-compartment mammillary model describing the disposition of propofol to find an optimal drug dosing control policy to drive the patient to a desired MRSS score.

5.2. Notation and Mathematical Preliminaries

In this section, we introduce notation, several definitions, and some key results concerning nonlinear nonnegative dynamical systems [61] that are necessary for developing the main results of this chapter. Specifically, for $x \in \mathbb{R}^n$ we write $x \geq\geq 0$ (resp., $x >> 0$) to indicate that every component of x is nonnegative (resp., positive). In this case, we say that x is *nonnegative* or *positive*, respectively. Likewise,

$A \in \mathbb{R}^{n\times m}$ is *nonnegative*[1] or *positive* if every entry of A is nonnegative or positive, respectively, which is written as $A \geq\geq 0$ or $A >> 0$, respectively. In addition, $\overline{\mathbb{R}}^n_+$ and $\mathbb{R}^n_+$ denote the nonnegative and positive orthants of $\mathbb{R}^n$, that is, if $x \in \mathbb{R}^n$, then $x \in \overline{\mathbb{R}}^n_+$ and $x \in \mathbb{R}^n_+$ are equivalent, respectively, to $x \geq\geq 0$ and $x >> 0$. Finally, we write $(\cdot)^{\mathrm{T}}$ to denote transpose, $\|\cdot\|$ for a vector norm in $\mathbb{R}^n$, $\mathbb{Z}$ to denote the set of integers, $\operatorname{dist}(x, \mathcal{M})$ to denote the distance of a point $x \in \mathbb{R}^n$ to the set $\mathcal{M} \subseteq \mathbb{R}^n$ in the norm $\|\cdot\|$ (that is, $\operatorname{dist}(x, \mathcal{M}) \triangleq \inf_{p\in\mathcal{M}} \|x-p\|$), and $\mathbf{e}$ to denote the ones vector of order n, that is, $\mathbf{e} \triangleq [1, \ldots, 1]^{\mathrm{T}}$.

The following definition introduces the notion of a nonnegative (resp., positive) function.

Definition 5.1. Let $T > 0$. A real function $u : [0, T] \to \mathbb{R}^m$ is a *nonnegative* (resp., *positive*) *function* if $u(t) \geq\geq 0$ (resp., $u(t) >> 0$) on the interval $[0, T]$.

The following definition introduces the notion of essentially nonnegative and compartmental vector fields [61].

Definition 5.2. Let $f = [f_1, \ldots, f_n]^{\mathrm{T}} : \mathcal{D} \subseteq \overline{\mathbb{R}}^n_+ \to \mathbb{R}^n$. Then f is *essentially nonnegative* if $f_i(x) \geq 0$, for all $i = 1, \ldots, n$, and $x \in \overline{\mathbb{R}}^n_+$ such that $x_i = 0$, where x_i denotes the ith component of x. f is *compartmental* if f is essentially nonnegative and $\mathbf{e}^{\mathrm{T}} f(x) \leq 0$, $x \in \overline{\mathbb{R}}^n_+$.

Note that if $f(x) = Ax$, where $A \in \mathbb{R}^{n\times n}$, then f is essentially nonnegative if and only if A is essentially nonnegative, that is, $A_{(i,j)} \geq 0$, $i, j = 1, ..., n$, $i \neq j$, where $A_{(i,j)}$ denotes the (i, j)th entry of A. Similarly, f is compartmental if and only if A is essentially nonnegative and $\sum_{i=1}^n A_{(i,j)} \leq 0$, $j = 1, \ldots, n$.

[1] In this dissertation, it is important to distinguish between a square nonnegative (resp., positive) matrix and a nonnegative-definite (resp., positive-definite) matrix.

In this chapter, we consider controlled nonlinear dynamical systems of the form

$$\dot{x}(t) = f(x(t)) + G(x(t))u(t), \quad x(0) = x_0, \quad t \geq 0, \tag{5.1}$$

where $x(t) \in \mathbb{R}^n$, $t \geq 0$, $u(t) \in \mathbb{R}^m$, $t \geq 0$, $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz continuous and satisfies $f(0) = 0$, $G : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ is continuous, and $u : [0, \infty) \to \mathbb{R}^m$ is piecewise continuous.

The following definition and proposition are needed for the main results of the chapter.

Definition 5.3. The nonlinear dynamical system given by (5.1) is *nonnegative* if, for every $x(0) \in \overline{\mathbb{R}}_+^n$ and $u(t) \geq\geq 0$, $t \geq 0$, the solution $x(t)$, $t \geq 0$, to (5.1) is nonnegative.

Proposition 5.1 [61]. The nonlinear dynamical system given by (5.1) is nonnegative if $f : \mathbb{R}^n \to \mathbb{R}^n$ is essentially nonnegative and $G(x) \geq\geq 0$, $x \in \overline{\mathbb{R}}_+^n$.

It follows from Proposition 5.1 that if $f(\cdot)$ is essentially nonnegative, then a nonnegative input signal $G(x(t))u(t)$, $t \geq 0$, is sufficient to guarantee the nonnegativity of the state of (5.1).

Finally, the following theorem and definition are needed for the main results of the chapter.

Theorem 5.1 [19]. Let $x \in \mathbb{R}^n$, $\mathcal{M} \subseteq \mathbb{R}^n$ be a closed set, and $\|\cdot\|$ be a norm in $\mathbb{R}^n$. Then there exists $a_x \in \mathcal{M}$ such that $\|x - a_x\| = \operatorname{dist}(x, \mathcal{M})$. Furthermore, if $\mathcal{M}$ is closed and convex, and $\|\cdot\| : \mathbb{R}^n \to \overline{\mathbb{R}}_+$ is strictly convex, then a_x is unique.

Definition 5.4. Let $x \in \mathbb{R}^n$, $\mathcal{M} \subseteq \mathbb{R}^n$ be a closed set, and $\|\cdot\|$ be a norm in $\mathbb{R}^n$. The *projection of x on $\mathcal{M}$* is given by

$$\operatorname{proj}_{\mathcal{M}}(x) \triangleq \{a \in \mathcal{M} : \|x - a\| = \operatorname{dist}(x, \mathcal{M})\}, \tag{5.2}$$

where $\text{proj}_{\mathcal{M}} : \mathbb{R}^n \to \mathbb{P}(\mathcal{M})$ and $\mathbb{P}(\mathcal{M})$ denotes the power set of $\mathcal{M}$.

Note that it follows from Theorem 5.1 and Definition 5.4 that if $p \in \text{proj}_{\mathcal{M}}(x)$, then $p = \text{argmin}_{a \in \mathcal{M}} \|x - a\|$. Finally, we note that if for every $x \in \mathbb{R}^n$ there exists a unique $p \in \text{proj}_{\mathcal{M}}(x)$, then $\mathcal{M}$ is closed and convex [19].

5.3. Nonlinear Compartmental Mammillary Systems

Drug dosing can be made more precise by using *pharmacokinetic* and *pharmacodynamic* modeling [55]. Pharmacokinetics is the study of the concentration of drugs in tissue as a function of time and dose schedule, while pharmacodynamics is the study of the relationship between drug concentration and drug effect. By relating dose to resultant drug concentration (pharmacokinetics) and concentration to effect (pharmacodynamics), a model for drug dosing can be generated.

Pharmacokinetic compartmental models typically assume that the body is comprised of multiple compartments. Within each compartment the drug concentration is assumed to be uniform due to perfect, instantaneous mixing. Transport to other compartments and elimination from the body occur by metabolic processes. For simplicity, the transport rate is often assumed to be proportional to drug concentration. Although the assumption of instantaneous mixing is an idealization, it has little effect on the accuracy of the model as long as we do not try to predict drug concentrations immediately after the initial drug dose.

In this section, we consider a nonlinear compartmental *mammillary* dynamical system to model the pharmacokinetics of a sedative drug. The nonlinear mammillary model is comprised of a central compartment from which there is outflow from the system and which exchanges material reversibly with one or more peripheral compartments. In an n-compartment mammillary model, the central compartment, which is

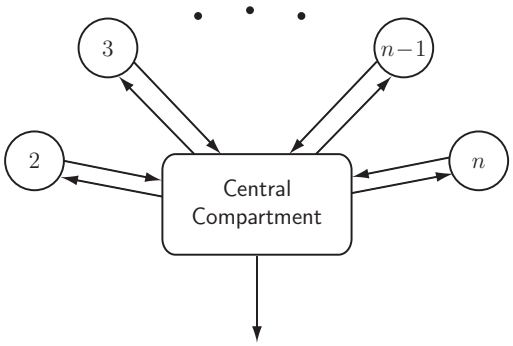

Figure 5.1: n-compartment mammillary model.

the site for drug administration, is generally thought to be comprised of the *intravascular blood* volume (i.e., blood within arteries and veins) as well as *highly perfused* organs (i.e., organs with high ratios of blood flow to weight) such as the heart, brain, kidneys, and liver. The central compartment exchanges drug with the peripheral compartments comprised of muscle, fat, and other organs and tissues of the body, which are metabolically inert as far as drug is concerned (see Figure 5.1).

The pharmacokinetic model of an n-compartment nonlinear mammillary model with a control input drug dose needed to achieve and maintain a target drug concentration is given by

$$\dot{x}_1(t) = -\left(\sum_{j=1}^{n} a_{j1}(c(t))\right) x_1(t) + \sum_{j=2}^{n} a_{1j}(c(t))x_j(t) + u(t), \quad x_1(0) = x_{10}, \quad t \geq 0, \tag{5.3}$$

$$\dot{x}_i(t) = a_{i1}(c(t))x_1(t) - a_{1i}(c(t))x_i(t), \quad x_i(0) = x_{i0}, \quad i = 2, \ldots, n, \tag{5.4}$$

where $c(t) = x_1(t)/V_{\mathrm{c}}$, V_{c} is the volume of the central compartment (about 15 l for a 70 kg patient), $a_{ij}(c)$, $i \neq j$, is the rate of transfer of drug from the jth compartment to the ith compartment, $a_{11}(c)$ is the rate of drug metabolism and elimination (metabolism typically occurs in the liver), and $u(t)$, $t \geq 0$, is the infusion rate of the sedative drug into the central compartment.

Although the concentration of the sedative agent in the blood is correlated with

lack of responsiveness [92], the concentration cannot be measured in real time. Since we are more interested in drug effect rather than drug concentration, we consider a model involving pharmacokinetics and pharmacodynamics for controlling consciousness. We use the sedation score to access the effect of anesthetic compounds on the brain. In Section 5.6, we utilize the modified probabilistic Hill equation [4] to model the relationship between the sedation score and the effect-site concentration. The effect-site compartment concentration is related to the concentration in the central compartment by the first-order model [127]

$$\dot{c}_{\text{eff}}(t) = a_{\text{eff}}(c(t) - c_{\text{eff}}(t)), \quad c_{\text{eff}}(0) = c(0), \quad t \geq 0, \tag{5.5}$$

where a_{eff} in min^{-1} is a positive time constant. In reality, the effect-site compartment equilibrates with the central compartment in a few minutes.

5.4. Hybrid Pharmacokinetic-Pharmacodynamic Model and Optimal Drug Dosing Policy

In this section, we model the pharmacokinetics and pharmacodynamics of a sedative agent as a hybrid deterministic-stochastic model involving the deterministic pharmacokinetic model developed in Section 5.3, and a stochastic pharmacodynamic model. Next, we use this model to develop an open-loop optimal drug dosing control policy for ICU sedation.

To develop our optimal control policy for ICU sedation, we rewrite the pharmacokinetic system (5.3), (5.4), and (5.5) as

$$\dot{x}(t) = f(x(t)) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \tag{5.6}$$

where $x = [x_1, \ldots, x_n, c_{\text{eff}}]^{\text{T}}$, $B = [1, 0_{1\times n}]^{\text{T}}$, and

$$f(x) \triangleq \begin{bmatrix} -\left(\sum_{j=1}^{n} a_{j1}(c)\right) x_1 + \sum_{j=2}^{n} a_{1j}(c)x_j \\ a_{21}(c)x_1 - a_{12}(c)x_2 \\ \vdots \\ a_{n1}(c)x_1 - a_{1n}(c)x_n \\ a_{\text{eff}}(c - c_{\text{eff}}) \end{bmatrix}. \tag{5.7}$$

Next, let the output $y(t)$ of the dynamical system (5.6) be given by a stochastic process. Specifically, for every $t \geq 0$, $y(t) = S(t)$ is a random variable with range$(S(t)) = \mathcal{S}$, where $\mathcal{S} \triangleq \{1, \ldots, 6\}$. Let the first-order distribution of the stochastic process $S(t)$, $t \geq 0$, be given by $F_S(s, c_{\text{eff}}) = P(S(t) \leq s)$, where $s \in \mathbb{R}$, $F_S : \mathcal{S} \times \mathcal{C} \to \mathbb{R}$, and $\mathcal{C} \subset \overline{\mathbb{R}}_+$ is a set of feasible drug concentrations in the effect-site compartment. The first-order distribution $F_S(s, c_{\text{eff}})$ is identified using experiments and statistical techniques, and provides a probabilistic relationship between the effect-site drug concentration c_{eff} and the sedation score. Finally, define the mapping $F : \mathcal{C} \to \mathbb{R}^6$ by

$$F(c_{\text{eff}}) \triangleq [F_S(1, c_{\text{eff}}), \ldots, F_S(6, c_{\text{eff}})]^{\text{T}}. \tag{5.8}$$

Our proposed approach for optimal drug dosing consists of two stages. In the first stage, the set of appropriate values of the drug concentration in the effect-site compartment denoted by $\mathcal{C}^*$ is identified such that the resulting probability distributions have desirable properties. More specifically, it is desirable to increase the probability associated with a desired sedation score (e.g., MRSS score of 3) and decrease the probabilities associated with all other levels of sedation. Ideally, we would like to target a cumulative distribution function for $S(t)$, $t > 0$, given by

$$F_{\text{step},S}(s) = \begin{cases} 0, & s < 3, \\ 1, & s \geq 3, \end{cases}, \tag{5.9}$$

where $s \in \mathbb{R}$. Define $F_{\text{step}} \triangleq [0, 0, 1, 1, 1, 1]^{\text{T}}$ and note that in general $F_{\text{step}} \notin \mathcal{F}$, where $\mathcal{F} \triangleq F(\mathcal{C})$ is the image of $\mathcal{C} \subset \overline{\mathbb{R}}_+$ under $F : \mathcal{C} \to \mathbb{R}^6$ defining the set of feasible

probability distributions given by

$$F(\mathcal{C}) \triangleq \{v : v = F(c) \text{ for some } c \in \mathcal{C}\}. \tag{5.10}$$

The following theorem and corollary provide a framework for identifying $\mathcal{C}^*$ given by

$$\mathcal{C}^* \triangleq F^{-1}\left(\text{proj}_{\mathcal{F}}(F_{\text{step}})\right), \tag{5.11}$$

where $F^{-1}(\mathcal{B}) \triangleq \{c \in \mathcal{C} : F(c) \in \mathcal{B}\}$, $\mathcal{B} \subset \mathbb{R}^6$, and $\text{proj}_{\mathcal{F}}(F_{\text{step}})$ is the projection of F_{step} on $\mathcal{F}$.

Theorem 5.2. Assume that the set of feasible drug concentrations $\mathcal{C} \subset \overline{\mathbb{R}}_+$ is closed and the mapping $F : \mathcal{C} \to \mathbb{R}^6$ is continuous. Then, $\mathcal{C}^*$ given by (5.11) is not empty. Furthermore, if $\mathcal{F}$ is convex, F is one-to-one, and $\|\cdot\| : \mathbb{R}^6 \to \overline{\mathbb{R}}_+$ is strictly convex, then $\mathcal{C}^*$ is a singleton.

Proof. Since $\mathcal{C}$ is closed and F is continuous, $\mathcal{F}$ is closed. Furthermore, it follows from Theorem 5.1 that there exists $G \in \mathcal{F}$ such that $\|F_{\text{step}} - G\| = \text{dist}(F_{\text{step}}, \mathcal{F})$, and hence, $G \in \text{proj}_{\mathcal{F}}(F_{\text{step}})$. Since $G \in \mathcal{F}$, there exists $c^* \in \mathcal{C}$ such that $F(c^*) = G$, and hence, $c^* \in \mathcal{C}^*$, which proves that $\mathcal{C}^*$ is not empty. If $\mathcal{F}$ is convex and $\|\cdot\| : \mathbb{R}^6 \to \overline{\mathbb{R}}_+$ is strictly convex, then it follows from Theorem 5.1 that $\text{proj}_{\mathcal{F}}(F_{\text{step}}) = \{G\}$. Now, since F is one-to-one, $\mathcal{C}^* = \{c^*\}$. □

Corollary 5.1. Assume that the set of feasible drug concentrations $\mathcal{C} \subset \overline{\mathbb{R}}_+$ is closed and the mapping $F : \mathcal{C} \to \mathbb{R}^6$ is continuous. Then,

$$F^{-1}\left(\text{proj}_{\mathcal{F}}(F_{\text{step}})\right) = \{c^* \in \mathcal{C} : c^* = \text{argmin}_{c \in \mathcal{C}} \|F_{\text{step}} - F(c)\|\}. \tag{5.12}$$

Proof. '$\subseteq$'. Let $c^* \in F^{-1}\left(\text{proj}_{\mathcal{F}}(F_{\text{step}})\right)$. Then, it follows that $\|F_{\text{step}} - F(c^*)\| = \text{dist}(F_{\text{step}}, \mathcal{F})$, where $\mathcal{F} = F(\mathcal{C})$. Thus,

$$F(c^*) = \text{argmin}_{F \in \mathcal{F}} \|F_{\text{step}} - F\| = \text{argmin}_{c \in \mathcal{C}} \|F_{\text{step}} - F(c)\|, \tag{5.13}$$

and hence, $c^* \in \{c^* \in \mathcal{C} : c^* = \text{argmin}_{c\in\mathcal{C}} \|F_{\text{step}} - F(c)\|\}$, which proves '$\subseteq$'.

'$\supseteq$'. Let $c^* = \text{argmin}_{c\in\mathcal{C}} \|F_{\text{step}} - F(c)\|$. Then, it follows that

$$\|F_{\text{step}} - F(c^*)\| = \min_{c\in\mathcal{C}} \|F_{\text{step}} - F(c)\| = \min_{F\in\mathcal{F}} \|F_{\text{step}} - F\| = \text{dist}(F_{\text{step}}, \mathcal{F}), \quad (5.14)$$

and hence, $c^* \in F^{-1}\left(\text{proj}_{\mathcal{F}}(F_{\text{step}})\right)$, which proves '$\supseteq$'. □

Note that once $\mathcal{C}^*$ is identified, an element of $\mathcal{C}^*$, denoted by c^*_{eff}, can be selected. The selected value $c^*_{\text{eff}} \in \mathcal{C}^*$ serves as the target drug concentration in the effect-site compartment. Using Corollary 5.1, c^*_{eff} can be identified by solving the optimization problem

$$\min_{c\in\mathcal{C}} \|F_{\text{step}} - F(c)\|. \quad (5.15)$$

Note that since it is desirable to reduce the probabilities associated with undersedation and oversedation, a specific norm can be used which enforces these properties. Specifically, we can choose the norm $\|\cdot\|_Q$, where $Q \in \mathbb{R}^{6\times 6}$ is a positive-definite weighting matrix and $\|z\|_Q^2 \triangleq z^{\mathrm{T}} Q z$, $z \in \mathbb{R}^6$. The weighting matrix Q can be used to assign weights (penalty) to different sedation levels. In particular, larger weighting values are assigned to sedation scores associated with undersedation and oversedation.

The second-stage of the proposed optimal drug dosing policy involves an open-loop optimal control problem whose solution is given by the following theorem.

Theorem 5.3. Consider the pharmacokinetic model (5.6) with initial condition $x_0 = [x_{10}, \ldots, x_{n0}, c_{\text{eff},0}]^{\mathrm{T}}$. Let the optimal sedative drug infusion rate $u^*(t)$, $t \geq 0$, be given by the solution to the minimization problem

$$\min_{u(\cdot)\in\mathcal{U}} \int_0^T L(x(t), u(t))\mathrm{d}t, \quad (5.16)$$

subject to

$$g(x, u) \leq\leq 0, \quad x \in \mathbb{R}^{n+1}, \quad u \in \mathbb{R}, \quad (5.17)$$

$$c_{\text{eff}}(T) \;=\; c^*_{\text{eff}}, \quad (5.18)$$

where

$$L(x,u) \triangleq \|F(c_{\rm eff}) - F(c^*_{\rm eff})\|^2_{R_1} + \frac{1}{2} r_2 u^2, \tag{5.19}$$

$$g(x,u) \triangleq [g_1(x,u),\, g_2(u)]^{\rm T}, \tag{5.20}$$

$$g_1(x,u) \triangleq (c_{\rm eff} - c_{\max})u, \tag{5.21}$$

$$g_2(u) \triangleq -u, \tag{5.22}$$

$c^*_{\rm eff} \in \mathcal{C}^*$, $\mathcal{C}^*$ is given by (5.11), $\mathcal{U} = \{u : [0,T] \to \mathbb{R} :\ u(\cdot)$ is piecewise continuous$\}$, $R_1 \in \mathbb{R}^{6\times 6}$ is a given positive-definite matrix, and $r_2 > 0$ and $c_{\max} > 0$ are given scalars. Then $u^*(t)$, $t \geq 0$, is given by

$$u^*(t) = \frac{1}{r_2}\left[-\lambda_1(t) - (c_{\rm eff} - c_{\max})\mu(t) + \nu(t)\right], \tag{5.23}$$

where $\lambda_1(t)$, $\mu(t)$, and $\nu(t)$, $t \geq 0$, are the solutions to

$$\begin{aligned}\dot{\lambda}_1(t) = {} & \left\{\left(\sum_{j=1}^{n} \frac{\partial a_{j1}(c)}{\partial c}\right)\frac{x_1(t)}{V_{\rm c}} + \sum_{j=1}^{n} a_{j1}(c) - \sum_{j=2}^{n} \frac{\partial a_{1j}(c)}{\partial c}\frac{x_j(t)}{V_{\rm c}}\right\}\lambda_1(t) \\ & - \sum_{j=2}^{n}\left(\frac{\partial a_{j1}(c)}{\partial c}\frac{x_1(t)}{V_{\rm c}} + a_{j1}(c) - \frac{\partial a_{1j}(c)}{\partial c}\frac{x_j(t)}{V_{\rm c}}\right)\lambda_j(t) - \frac{a_{\rm eff}}{V_{\rm c}}\lambda_{n+1}, \end{aligned} \tag{5.24}$$

$$\dot{\lambda}_i(t) = -a_{1i}(c)\left[\lambda_1(t) - \lambda_i(t)\right], \quad i = 2,\ldots,n, \tag{5.25}$$

$$\begin{aligned}\dot{\lambda}_{n+1}(t) = {} & a_{\rm eff}\lambda_{n+1}(t) - 2\left[F(c_{\rm eff}(t)) - F(c^*_{\rm eff})\right]^{\rm T} R_1 \frac{\partial F(c_{\rm eff})}{\partial c_{\rm eff}} \\ & -\mu(t)\frac{1}{r_2}\left[-\lambda_1(t) - (c_{\rm eff} - c_{\max})\mu(t) + \nu(t)\right], \end{aligned} \tag{5.26}$$

with boundary conditions

$$x(0) = x_0, \tag{5.27}$$

$$c_{\rm eff}(T) = c^*_{\rm eff}, \tag{5.28}$$

$$\lambda_i(T) = 0, \quad i = 1,\ldots,n, \tag{5.29}$$

and $x(t)$, $t \geq 0$, satisfying (5.6), $\mu(t) \geq 0$, $t \geq 0$, if $g_1(x(t),u(t)) = 0$, $\mu(t) = 0$ if $g_1(x(t),u(t)) < 0$, $\nu(t) \geq 0$, $t \geq 0$, if $g_2(u(t)) = 0$, and $\nu(t) = 0$ if $g_2(u(t)) < 0$, $t \geq 0$. Furthermore, $u^*(t) \geq 0$, $t \geq 0$, and $x(t) \geq\geq 0$, $t \geq 0$, for all $x_0 \in \overline{\mathbb{R}}^{n+1}_+$.

Proof. Equations (5.23)–(5.26) are a direct consequence of the first-order necessary conditions for optimality of the optimization problem (5.16)–(5.18). Now, since $g(x,u) \leq\leq 0$, $(x,u) \in \mathbb{R}^{n+1} \times \mathbb{R}$, it follows that $u^*(t) \geq 0$, $t \geq 0$. Finally, since $f(x)$ given by (5.7) is essentially nonnegative, it follows from Proposition 5.1 that $x(t) \geq\geq 0$, $t \geq 0$, for all $x_0 \in \overline{\mathbb{R}}_+^{n+1}$. □

Note that the cost functional given by (5.16) penalizes the control effort as well as the deviations from the cumulative distribution function $F(c^*_{\text{eff}})$. In addition, the inequality constraint (5.17) ensures that the control input $u(t)$, $t \geq 0$, is nonnegative and the drug concentration in the effect-site compartment does not exceed the maximum concentration $c_{\max}$. Furthermore, the equality constraint (5.18) ensures that the drug concentration in the effect-site compartment reaches the target drug concentration c^*_{eff} in finite time T. Finally, note that since $r_2 > 0$, it follows from the *Legendre-Clebsch condition* [25] that $u^*(t)$, $t \geq 0$, given by (5.23) is a minimizer.

5.5. Nonlinear Pharmacokinetic Model for Disposition of Propofol

In this section, we use nonnegative and compartmental modeling to model the pharmacokinetics of the sedative agent propofol. Propofol, or *2,6-diisopropylphenol*, is an intravenous hypnotic agent that in low doses can produce anxiolysis and in higher doses hypnosis (i.e., lack of responsiveness and lack of consciousness). Propofol is widely used for ICU sedation because of this spectrum of pharmacodynamic effects and also because of its pharmacokinetics. It is typically administered as a continuous infusion and it is a short acting drug that can be readily titrated, that is, if the infusion rate is increased the blood level increases relatively quickly. Hence, the pharmacological effect of the drug can be quickly varied by varying the infusion rate.

Propofol is a *myocardial* depressant, that is, it decreases the contractility of the

heart and lowers *cardiac output* (i.e., the volume of blood pumped by the heart per unit time). As a consequence, decreased cardiac output slows down redistribution kinetics, that is, the transfer of blood from the central compartment (heart, brain, kidneys, and liver) to the peripheral compartments (muscle and fat). In addition, decreased cardiac output could increase drug concentrations in the central compartment, causing even more myocardial depression and further decrease in cardiac output. This instability can lead to oversedation.

Oversedation increases risk to the patient since liberation from mechanical ventilation, one of the most common life-saving procedures performed in the ICU, may not be possible due to a diminished level of consciousness and respiratory depression from sedative drugs resulting in prolonged length of stay in the ICU. Prolonged ventilation is expensive and is associated with known risks, such as inadvertent extubation, laryngo-tracheal trauma, and ventilator-associated pneumonia. Alternatively, undersedation leads to agitation and can result in dangerous situations for both the patient and the intensivist. Specifically, agitated patients can do physical harm to themselves by dislodging their endotracheal tube which can potentially endanger their life.

The pharmacokinetics of propofol are described by the three-compartment model [61, 103] shown in Figure 5.2, where x_1 denotes the mass of drug in the central compartment, which, as discussed in Section 5.3, is the site for drug administration and is generally thought to be comprised of the intravascular blood volume as well as highly perfused organs such as the heart, brain, kidneys, and liver. These organs receive a large fraction of the cardiac output. The remainder of the drug in the body is assumed to reside in two peripheral compartments, one identified with muscle and one with fat; the masses in these compartments are denoted by x_2 and x_3, respectively. These compartments receive less than 20% of the cardiac output.

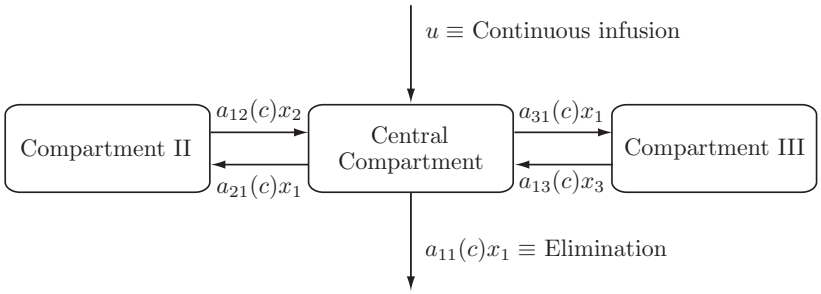

Figure 5.2: Pharmacokinetic model for disposition of propofol.

A mass balance of the three-state compartmental model yields

$$\begin{aligned}\dot{x}_1(t) = & -[a_{11}(c(t)) + a_{21}(c(t)) + a_{31}(c(t))]x_1(t) \\ & +a_{12}(c(t))x_2(t) + a_{13}(c(t))x_3(t) + u(t), \quad x_1(0) = x_{10}, \quad t \geq 0, \end{aligned} \tag{5.30}$$

$$\dot{x}_2(t) = a_{21}(c(t))x_1(t) - a_{12}(c(t))x_2(t), \quad x_2(0) = x_{20}, \tag{5.31}$$

$$\dot{x}_3(t) = a_{31}(c(t))x_1(t) - a_{13}(c(t))x_3(t), \quad x_3(0) = x_{30}, \tag{5.32}$$

where $c(t) = x_1(t)/V_{\mathrm{c}}$, V_{c} is the volume of the central compartment (about 15 l for a 70 kg patient), $a_{ij}(c)$, $i \neq j$, is the rate of transfer of drug from the jth compartment to the ith compartment, $a_{11}(c)$ is the rate of drug metabolism and elimination (metabolism typically occurs in the liver), and $u(t)$, $t \geq 0$, is the infusion rate of the sedative drug propofol into the central compartment. The transfer coefficients are assumed to be functions of the drug concentration c since it is well known that the pharmacokinetics of propofol are influenced by cardiac output [143] and, in turn, cardiac output is influenced by propofol plasma concentrations, both due to *venodilation* (pooling of blood in dilated vains) [110] and myocardial depression [79].

Experimental data indicate that the transfer coefficients $a_{ij}(\cdot)$ are nonincreasing functions of the propofol concentration [79, 110]. The most widely used empirical models for pharmacodynamic concentration-effect relationships are modifications of the Hill equation [73]. Applying this almost ubiquitous empirical model to the rela-

tionship between transfer coefficients implies that

$$a_{ij}(c) = A_{ij}Q_{ij}(c), \quad Q_{ij}(c) = Q_0\tilde{C}_{50,ij}^{\alpha_{ij}}/(\tilde{C}_{50,ij}^{\alpha_{ij}} + c^{\alpha_{ij}}), \tag{5.33}$$

where, for $i, j \in \{1, 2, 3\}$, $i \neq j$, $\tilde{C}_{50,ij}$ is the drug concentration associated with a 50% decrease in the transfer coefficient, α_{ij} is a parameter that determines the steepness of the concentration-effect relationship, and A_{ij} are positive constants. Note that both pharmacokinetic parameters are functions of i and j, that is, there are distinct Hill equations for each transfer coefficient. Furthermore, since for many drugs the rate of metabolism $a_{11}(c)$ is proportional to the rate of transport of drug to the liver we assume that $a_{11}(c)$ is also proportional to the cardiac output so that $a_{11}(c) = A_{11}Q_{11}(c)$. Finally, the relationship between the effect-site and the central compartment is given by (5.5).

5.6. Optimal Drug Dosing Policy for Propofol

The framework presented in Section 5.4 is applicable to sedative agents for which a valid compartmental model capturing the pharmacokinetics and an associated probabilistic model capturing drug concentration and sedation score exist. In this section, we use the framework developed in Section 5.4 to model the pharmacokinetics and pharmacodynamics of propofol as a hybrid deterministic-stochastic model. Specifically, we use the deterministic pharmacokinetic model developed in Section 5.5. Next, we use this model to develop an open-loop optimal drug dosing control policy for ICU sedation.

In [4] the authors investigate the relationship between drug concentration and the ICU patient's sedation score. Specifically, the sedation score is modeled as a random variable and an empirical cumulative distribution function for this random variable is developed and validated for propofol-based sedation where the cumulative

distribution function is a function of drug concentration at the effect site.

To develop our optimal control policy for ICU sedation, we rewrite the pharmacokinetic system (5.5), (5.30)–(5.32) as

$$\dot{x}(t) = f(x(t)) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \tag{5.34}$$

where $x = [x_1,\, x_2,\, x_3,\, c_{\text{eff}}]^{\text{T}}$, $B = [1,\, 0,\, 0,\, 0]^{\text{T}}$, and

$$f(x) \triangleq \begin{bmatrix} -\left(a_{11}(c) + a_{21}(c) + a_{31}(c)\right) x_1 + a_{12}(c)x_2 + a_{13}(c)x_3 \\ a_{21}(c)x_1 - a_{12}(c)x_2 \\ a_{31}(c)x_1 - a_{13}(c)x_3 \\ a_{\text{eff}}(c - c_{\text{eff}}) \end{bmatrix}. \tag{5.35}$$

Next, let the output $y(t)$ of the dynamical system (5.34) be given by a stochastic process. Specifically, for every $t \geq 0$, $y(t) = S(t)$ is a random variable with range$(S(t)) = \mathcal{S}$, where $\mathcal{S} \triangleq \{1, \ldots, 6\}$. The first-order distribution of the stochastic process $S(t)$ is given by [4]

$$F_S(s, c_{\text{eff}}) = P(S(t) \leq s) = \begin{cases} 0 & s < 1, \\ 1 - \frac{c_{\text{eff}}^{\gamma}(t)}{c_{\text{eff}}^{\gamma}(t) + C_{50,\lfloor s \rfloor + 1}^{\gamma}}, & 1 \leq s < 6, \\ 1, & s \geq 6, \end{cases} \tag{5.36}$$

where $s \in \mathbb{R}$, $F_S : \mathcal{S} \times \mathcal{C} \to \mathbb{R}$ is a first-order distribution function of the stochastic process $S(t)$, $\mathcal{C} \subset \overline{\mathbb{R}}_+$ is a closed set of feasible drug concentrations in the effect-site compartment, $\lfloor \cdot \rfloor$ denotes the floor function defined by $\lfloor s \rfloor \triangleq \max_{z \in \mathbb{Z}} z \leq s$, and $\gamma > 0$ is a factor determining the steepness of the concentration-effect relationship. Finally, note that $F : \mathcal{C} \to \mathbb{R}^6$ given by (5.8) is continuous.

The second-stage of the proposed optimal drug dosing policy involves an open-loop optimal control problem. Specifically, it follows from Theorem 5.3 that the optimal propofol infusion rate $u^*(t)$, $t \geq 0$, is given by the solution to the minimization problem (5.16) subject to (5.17)–(5.18), where $c_{\text{eff}}^* \in \mathcal{C}^*$, and $\mathcal{C}^*$ is given by (5.11). In particular, $u^*(t)$, $t \geq 0$, is given by

$$u^*(t) = \frac{1}{r_2}\left[-\lambda_1(t) - (c_{\text{eff}} - c_{\max})\mu(t) + \nu(t)\right], \tag{5.37}$$

where $\lambda_1(t)$, $\mu(t)$, and $\nu(t)$, $t \geq 0$, are the solutions to

$$\begin{aligned}\dot{\lambda}_1(t) = &\left[\left(\frac{\partial a_{11}(c)}{\partial c} + \frac{\partial a_{21}(c)}{\partial c} + \frac{\partial a_{31}(c)}{\partial c}\right)\frac{x_1(t)}{V_c} + a_{11}(c) + a_{21}(c) + a_{31}(c)\right.\\ &\left. - \frac{\partial a_{12}(c)}{\partial c}\frac{x_2(t)}{V_c} - \frac{\partial a_{13}(c)}{\partial c}\frac{x_3(t)}{V_c}\right]\lambda_1(t) + \left[-\frac{\partial a_{21}(c)}{\partial c}\frac{x_1(t)}{V_c} - a_{21}(c)\right.\\ &\left. + \frac{\partial a_{12}(c)}{\partial c}\frac{x_2(t)}{V_c}\right]\lambda_2(t) + \left[-\frac{\partial a_{31}(c)}{\partial c}\frac{x_1(t)}{V_c} - a_{31}(c) + \frac{\partial a_{13}(c)}{\partial c}\frac{x_3(t)}{V_c}\right]\lambda_3(t)\\ &- \frac{a_{\text{eff}}}{V_c}\lambda_4(t), \end{aligned} \tag{5.38}$$

$$\dot{\lambda}_2(t) = -a_{12}(c)\lambda_1(t) + a_{12}\lambda_2(t), \tag{5.39}$$

$$\dot{\lambda}_3(t) = -a_{13}(c)\lambda_1(t) + a_{13}(c)\lambda_3(t), \tag{5.40}$$

$$\begin{aligned}\dot{\lambda}_4(t) = &\, a_{\text{eff}}\lambda_4(t) - 2\left[F(c_{\text{eff}}(t)) - F(c^*_{\text{eff}})\right]^{\mathrm{T}} R_1 \frac{\partial F(c_{\text{eff}})}{\partial c_{\text{eff}}}\\ &- \mu(t)\frac{1}{r_2}\left[-\lambda_1(t) - (c_{\text{eff}} - c_{\max})\mu(t) + \nu(t)\right], \end{aligned} \tag{5.41}$$

where

$$\frac{\partial a_{ij}(c)}{\partial c} = \frac{-\alpha_{ij}c^{\alpha_{ij}-1}A_{ij}Q_0\tilde{C}_{50,ij}^{\alpha_{ij}}}{(\tilde{C}_{50,ij}^{\alpha_{ij}} + c^{\alpha_{ij}})^2}, \quad i = 1,\ j = 1, \text{ and } i, j \in \{1, 2, 3\},\ i \neq j, \tag{5.42}$$

$$\frac{\partial F(c_{\text{eff}})}{\partial c_{\text{eff}}} = \left[-\frac{\gamma c_{\text{eff}}^{\gamma-1}C_{50,2}^{\gamma}}{(c_{\text{eff}}^{\gamma}+C_{50,2}^{\gamma})^2}, -\frac{\gamma c_{\text{eff}}^{\gamma-1}C_{50,3}^{\gamma}}{(c_{\text{eff}}^{\gamma}+C_{50,3}^{\gamma})^2}, -\frac{\gamma c_{\text{eff}}^{\gamma-1}C_{50,4}^{\gamma}}{(c_{\text{eff}}^{\gamma}+C_{50,4}^{\gamma})^2}, -\frac{\gamma c_{\text{eff}}^{\gamma-1}C_{50,5}^{\gamma}}{(c_{\text{eff}}^{\gamma}+C_{50,5}^{\gamma})^2}, -\frac{\gamma c_{\text{eff}}^{\gamma-1}C_{50,6}^{\gamma}}{(c_{\text{eff}}^{\gamma}+C_{50,6}^{\gamma})^2}, 0\right]^{\mathrm{T}}, \tag{5.43}$$

with boundary conditions (5.27), (5.28), and $\lambda_1(T) = \lambda_2(T) = \lambda_3(T) = 0$, and $x(t)$, $t \geq 0$, satisfying (5.34), $\mu(t) \geq 0$, $t \geq 0$, if $g_1(x(t), u(t)) = 0$, $\mu(t) = 0$ if $g_1(x(t), u(t)) < 0$, $\nu(t) \geq 0$, $t \geq 0$, if $g_2(u(t)) = 0$, and $\nu(t) = 0$ if $g_2(u(t)) < 0$, $t \geq 0$. Furthermore, $u^*(t) \geq 0$, $t \geq 0$, and $x(t) \geq\geq 0$, $t \geq 0$, for all $x_0 \in \overline{\mathbb{R}}^4_+$.

Remark 5.1. The framework in this chapter can be used for other sedative agents for which a valid compartmental model capturing the pharmacokinetics and an associated probabilistic model capturing drug concentration and sedation score exist. For example, the pharmacokinetics of *midazolam* (an alternative intravenous sedative agent used as a hypnotic) is described by a two-compartment model [67]. The

empirical relationship between drug concentration and sedation score for midazolam is developed in [5]. Using an identical procedure as outlined above, the optimal drug dosing policy for the midazolam infusion rate can be found.

5.7. Illustrative Numerical Example

In this section, we present a numerical example to demonstrate the efficacy of the proposed framework. For simplicity of exposition and to provide a nonlinear model to illustrate implementation of our open-loop optimal controller, we assume that $\tilde{C}_{50}$ and α in (5.33) are independent of i and j [61]. Furthermore, since decreases in cardiac output are observed at clinically utilized propofol concentrations, we arbitrarily assign $\tilde{C}_{50}$ a value of 4 μg/ml since this value is in the mid-range of clinically utilized values. We also arbitrarily assign α a value of 3 [93]. This value is within the typical range of those observed for ligand-receptor binding (see the discussion in [43]). Note that these assumptions on $\tilde{C}_{50}$ and α (both the independence from i and j and the assumed values) are done to provide a numerical framework for simulation. Even if these assumptions are incorrect, the basic Hill equations relating the transfer coefficients to propofol concentration are consistent with standard pharmacodynamic modeling.

For our simulation we assume $V_{\mathrm{c}} = (0.228\ \mathrm{l/kg})(M\ \mathrm{kg})$, where $M = 70$ kg is the mass of the patient, $A_{21}Q_0 = 0.112\ \mathrm{min}^{-1}$, $A_{12}Q_0 = 0.055\ \mathrm{min}^{-1}$, $A_{31}Q_0 = 0.0419\ \mathrm{min}^{-1}$, $A_{13}Q_0 = 0.0033\ \mathrm{min}^{-1}$, $A_{11}Q_0 = 0.119\ \mathrm{min}^{-1}$, $\alpha = 3$, and $\tilde{C}_{50} = 4\ \mu\mathrm{g/ml}$ [93, 103]. Note that the parameter values for α and $\tilde{C}_{50}$ probably exaggerate the effect of propofol on cardiac output. They have been selected to accentuate nonlinearity but they are not biologically unrealistic. Furthermore, in (5.36) we assume $C_{50,2} = 0.13\,\mu$g/ml, $C_{50,3} = 0.50\,\mu$g/ml, $C_{50,4} = 0.74\,\mu$g/ml, $C_{50,5} = 1.48\,\mu$g/ml, $C_{50,6} = 2.34\,\mu$g/ml, and $\gamma = 1.7$ [4]. In addition, we assume $T = 5$ min, $Q = R_1 = \mathrm{diag}[17, 2, 1, 2, 17, 82]$, and $r_2 = 0.01$. Using (5.15) the optimal effect-site

drug concentration was found to be $c^*_{\text{eff}} = 0.60294\,\mu\text{g/ml}$.

For our simulation we choose the diagonal matrix R_1 with diagonal entries given by $R_{1(i,i)} = (i-3)^4 + 1$, $i = 1, \ldots, 6$. This ensures that a larger weight (penalty) is assigned to sedation scores associated with undersedation and oversedation. The drug concentration of the central compartment and the effect-site compartment as well as control input as a function of time are shown in Figures 5.3 and 5.4, respectively. The probability mass function of the sedation score is given in Figure 5.5 for $t = 0$, 1, 3, and 5 min. Note that at $t = 5$ min the probability that the patient has an MRSS sedation score of 2, 3, or 4 (i.e., the patient is lightly sedated, moderately sedated and follows simple commands, or deeply sedated and responds to nonpainful stimuli) is 75%.

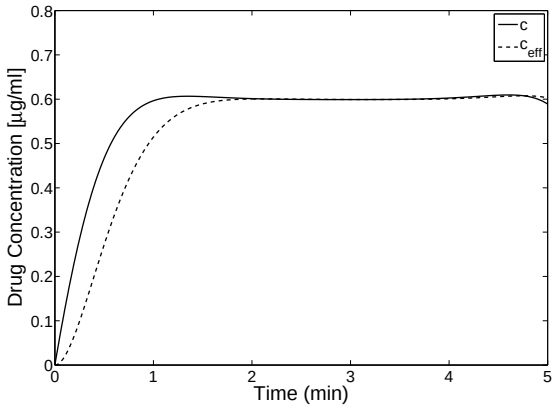

Figure 5.3: Drug concentration $c(t) = \frac{x_1(t)}{V_c}$ and $c_{\text{eff}}(t)$ as a function of time.

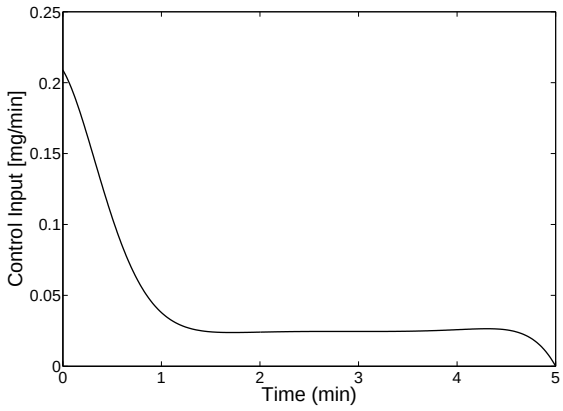

Figure 5.4: Control input as a function of time.

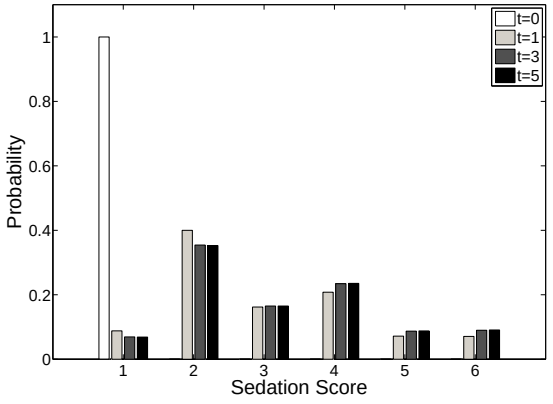

Figure 5.5: Probability mass function for sedation score $S(t)$ for $t = 0$, 1, 3, and 5.

Chapter 6

$\mathcal{H}_2$ and Mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ Suboptimal Estimation and Control for Nonnegative Dynamical Systems

6.1. Introduction

Nonnegative dynamical systems involve dynamic states whose values are nonnegative [48,62,89,91]. A subclass of nonnegative dynamical systems are compartmental systems [62,83,85]. Compartmental systems involve dynamical models that are characterized by conservation laws (e.g., mass, energy, fluid, etc.) capturing the exchange of material between coupled macroscopic subsystems known as compartments. These models are widespread in biological, physiological, and ecological sciences as well as engineering systems such as queuing, large-scale, telecommunications, transportation, power, and network systems, to cite but a few examples (see [48,62] and the references therein). Since nonnegative and compartmental systems have specialized structures, special control law strategies need to be developed that guarantee that the trajectories of the closed-loop plant system states remain in the nonnegative orthant of the state space for nonnegative initial conditions. In addition, for certain applications of nonnegative systems, such as active control for clinical pharmacology, we require the control (source) inputs to be nonnegative.

Even though nonnegative systems are often encountered in numerous application

areas, nonnegative orthant stabilizability and holdability has received little attention in the literature. Notable exceptions include [9, 98]. In addition, fixed-structure control for linear nonnegative dynamical systems, and adaptive and neuroadaptive control of nonnegative and compartmental systems have been recently developed in [63,68–70,113]. In this chapter, we use linear matrix inequalities (LMIs) to develop $\mathcal{H}_2$ and mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ (sub)optimal estimators and controllers for nonnegative dynamical systems. Linear matrix inequalities provide a powerful design framework for linear control problems [18, 42, 115, 125]. Since LMIs lead to convex or quasiconvex optimization problems, they can be solved very efficiently using interior-point algorithms. An intersting feature of nonnegative orthant stabilizability is that it can be formulated as a solution to an LMI problem. However, $\mathcal{H}_2$ optimal nonnegative orthant stabilizability cannot, in general, be formulated as an LMI problem. In this chapter, we formulate a series of generalized eigenvalue problems subject to a set of LMI constraints for designing $\mathcal{H}_2$ suboptimal estimators, static controllers, and dynamic controllers for nonnegative dynamical systems.

6.2. Notation and Mathematical Preliminaries

In this section, we introduce notation, several definitions, and some key results concerning linear nonnegative dynamical systems [9, 10, 13, 62] that are necessary for developing the main results of this chapter. Specifically, $\mathrm{spec}(A)$ denotes the spectrum of A, $(\cdot)^{\mathrm{T}}$ denotes transpose, $(\cdot)^{-1}$ denotes inverse, $\sigma_{\max}(\cdot)$ denotes maximum singular value, and $\mathbb{E}$ denotes expectation.

The following definition introduces the notion of essentially nonnegative, compartmental, and M-matrices.

Definition 6.1 [13, 62]. Let $A \in \mathbb{R}^{n\times n}$. A is *essentially nonnegative* if $A_{(i,j)} \geq 0$,

$i, j = 1, \ldots, n$, $i \neq j$. A is *compartmental* if A is essentially nonnegative and $A^{\mathrm{T}}\mathbf{e} \leq\leq 0$. A is an M-*matrix* (resp., *nonsingular* M-*matrix*) if $-A$ is essentially nonnegative and all the principal minors of A are nonnegative (resp., positive).

Next, consider the linear nonnegative dynamical system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.1}$$

where $x(t) \in \mathbb{R}^n$, $t \geq 0$, and $A \in \mathbb{R}^{n\times n}$ is essentially nonnegative. The solution to (6.1) is standard and is given by $x(t) = e^{At}x(0)$, $t \geq 0$. The following proposition proven in [13] (see also [62]) shows that A is essentially nonnegative if and only if the state transition matrix e^{At} is nonnegative on $[0, \infty)$.

Proposition 6.1 [13, 62]. Let $A \in \mathbb{R}^{n\times n}$. Then A is essentially nonnegative if and only if e^{At} is nonnegative for all $t \geq 0$. Furthermore, if A is essentially nonnegative and $x_0 \geq\geq 0$, then $x(t) \geq\geq 0$, $t \geq 0$, where $x(t)$, $t \geq 0$, denotes the solution to (6.1).

The following theorem gives necessary and sufficient conditions for asymptotic stability of a linear nonnegative dynamical system using a quadratic component decoupled Lyapunov function.

Theorem 6.1 [62]. Consider the linear dynamical system $\mathcal{G}$ given by (6.1) where $A \in \mathbb{R}^{n\times n}$ is essentially nonnegative. Then $\mathcal{G}$ is asymptotically stable if and only if there exist a positive diagonal matrix $P \in \mathbb{R}^{n\times n}$ and an $n \times n$ positive-definite matrix R such that

$$0 = A^{\mathrm{T}}P + PA + R. \tag{6.2}$$

Remark 6.1. Note that it follows from Theorem 6.1 that the linear dynamical system $\mathcal{G}$ is asymptotically stable if and only if there exists a positive diagonal matrix

$P \in \mathbb{R}^{n\times n}$ such that the following LMI holds

$$0 > A^{\mathrm{T}}P + PA. \tag{6.3}$$

Finally, in this chapter we consider controlled dynamical systems of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.4}$$

where $x(t) \in \mathbb{R}^n$, $t \geq 0$, $u(t) \in \mathbb{R}^m$, $t \geq 0$, $A \in \mathbb{R}^{n\times n}$, and $B \in \mathbb{R}^{n\times m}$. The following definition and proposition are needed for the main results of the chapter.

Definition 6.2 [62]. The linear dynamical system given by (6.4) is *nonnegative* if for every $x(0) \in \overline{\mathbb{R}}^n_+$ and $u(t) \geq\geq 0$, $t \geq 0$, the solution $x(t)$, $t \geq 0$, to (6.4) is nonnegative.

Proposition 6.2 [62]. The linear dynamical system given by (6.4) is nonnegative if and only if $A \in \mathbb{R}^{n\times n}$ is essentially nonnegative and $B \in \mathbb{R}^{n\times m}$ is nonnegative.

It follows from Proposition 6.2 that the weighted control input signal $Bu(t)$, $t \geq 0$, needs to be nonnegative to guarantee the nonnegativity of the state of (6.4). This is due to the fact that when the initial state of (6.4) belongs to the boundary of the nonnegative orthant, a negative input can destroy the nonnegativity of the state of (6.4). Alternatively, however, if the initial state is in the interior of the nonnegative orthant, then it follows from continuity of solutions with respect to system initial conditions that, over a small interval of time, nonnegativity of the state of (6.4) is guaranteed *irrespective* of the sign of each element of the weighted control input $Bu(t)$ over this time interval. However, unlike open-loop control wherein lack of coordination between the input and the state necessitates nonnegativity of the control input, a *feedback* control signal predicated on the system state variables allows for

the anticipation of loss of nonnegativity of the state. Hence, state feedback control signals can take negative values while ensuring nonnegativity of the system states. For further discussion of the above fact see [98].

6.3. $\mathcal{H}_2$ Suboptimal Control for Nonnegative Dynamical Systems

In this section, we present a suboptimal control framework for minimizing the $\mathcal{H}_2$ cost of a linear nonnegative dynamical system while constraining the system states to the nonnegative orthant of the state space. First, however, we recall the following standard results.

Theorem 6.2 [41]. Let $A \in \mathbb{R}^{n\times n}$. A is Hurwitz if and only if there exists an $n \times n$ matrix $P > 0$ such that $A^{\mathrm{T}}P + PA < 0$.

Next, consider the linear dynamical system $\mathcal{G}$ given by

$$\dot{x}(t) = Ax(t) + Dw(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.5}$$

$$z(t) = Ex(t), \tag{6.6}$$

where $A \in \mathbb{R}^{n\times n}$, $D \in \mathbb{R}^{n\times d}$, $E \in \mathbb{R}^{p\times n}$, and $w(t) \in \mathbb{R}^d$ is a standard white noise process.

Theorem 6.3 [41]. Consider the linear dynamical system $\mathcal{G}$ given by (6.5) and (6.6). Then A is Hurwitz and $|||G|||_2^2 < \lambda$ if and only if there exists $Q > 0$ such that $\operatorname{tr} EQE^{\mathrm{T}} < \lambda$ and

$$0 > AQ + QA^{\mathrm{T}} + DD^{\mathrm{T}}, \tag{6.7}$$

where $|||G|||_2$ denotes the $\mathcal{H}_2$ norm of the transfer function $G(s) = E(sI_n - A)^{-1}D$.

Remark 6.2. Recall that the $\mathcal{H}_2$ norm of $G(s)$ is given by

$$|||G|||_2^2 = \lim_{t\to\infty} \mathbb{E}\left\{\frac{1}{t}\int_0^t z^{\mathrm{T}}(s)z(s)\mathrm{d}s\right\}. \tag{6.8}$$

Next, we consider the static $\mathcal{H}_2$ optimal regulator problem and present the solution to this problem using LMIs.

Problem 6.1 (Static $\mathcal{H}_2$ Optimal Regulator Problem). Given the linear controlled system

$$\dot{x}(t) = Ax(t) + Bu(t) + Dw(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.9}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $D \in \mathbb{R}^{n\times d}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ is the control input, $w(t) \in \mathbb{R}^d$ is a standard white noise process, and (A, B) is stabilizable, determine a static state feedback control law $u(t) = Kx(t)$, where $K \in \mathbb{R}^{m\times n}$, that satisfies the following design criteria: *i)* the closed-loop system matrix given by $A + BK$ is Hurwitz; and *ii)* the $\mathcal{H}_2$ performance criterion

$$J(K) = \lim_{t\to\infty} \mathbb{E}\left\{\frac{1}{t}\int_0^t z^{\mathrm{T}}(s)z(s)\mathrm{d}s\right\} \tag{6.10}$$

is minimized, where the performance variable $z(t) \in \mathbb{R}^p$ is given by

$$z(t) = E_1x(t) + E_2u(t), \tag{6.11}$$

and where $E_1 \in \mathbb{R}^{p\times n}$ and $E_2 \in \mathbb{R}^{p\times m}$ are such that $E_1^{\mathrm{T}}E_2 = 0$.

Theorem 6.4 [41]. Consider the Static $\mathcal{H}_2$ Optimal Regulator Problem given by Problem 6.1. Then the optimal control gain and optimal $\mathcal{H}_2$ cost are given by $K^* = Z^*X^{*-1}$ and $J(K^*) = \operatorname{tr} W^*$, respectively, where $X^* = X^{*\mathrm{T}} \in \mathbb{R}^{n\times n}$, $W^* = W^{*\mathrm{T}} \in \mathbb{R}^{p\times p}$, and $Z^* \in \mathbb{R}^{m\times n}$ are the optimal solutions to the generalized eigenvalue problem (GEVP)

$$\inf_{X\in\mathbb{R}^{n\times n},\, Z\in\mathbb{R}^{m\times n},\, W\in\mathbb{R}^{p\times p}} \operatorname{tr} W \tag{6.12}$$

subject to

$$0 > AX + XA^{\mathrm{T}} + BZ + Z^{\mathrm{T}}B^{\mathrm{T}} + DD^{\mathrm{T}}, \tag{6.13}$$

$$0 < \begin{bmatrix} X & (E_1X + E_2Z)^{\mathrm{T}} \\ (E_1X + E_2Z) & W \end{bmatrix}. \tag{6.14}$$

It is important to note that the generalized eigenvalue problem formulated in Theorem 6.4 is a convex optimization problem, and hence, the infimum (6.12) is attained. Similar remarks hold for all the theorems in this chapter. Before stating the main theorem of this section we need the following definitions and proposition.

Definition 6.3 [9]. Consider the linear dynamical system given by (6.4). Then the pair (A, B) is *nonnegative orthant holdable* if there exists a control input $u : [\,0, \infty) \to \mathbb{R}^m$ such that, with initial condition $x_0 \in \overline{\mathbb{R}}^n_+$, $x(t) \in \overline{\mathbb{R}}^n_+$ for all $t \geq 0$.

Definition 6.4 [9]. Consider the linear dynamical system given by (6.4). Then the pair (A, B) is *nonnegative orthant feedback holdable* if there exists a feedback control law of the form $u(t) = Kx(t)$ such that, for every initial condition $x_0 \in \overline{\mathbb{R}}^n_+$, $x(t) \in \overline{\mathbb{R}}^n_+$ for all $t \geq 0$.

Recall that the linear dynamical system (6.4) is *stabilizable* if and only if there exists an $n \times n$ matrix $Y > 0$ such that

$$0 > (A + BK)^{\mathrm{T}}Y + Y(A + BK). \tag{6.15}$$

Note that (6.15) is not an LMI in K and Y. However, there exist an $n \times n$ matrix $Y > 0$ and a matrix $K \in \mathbb{R}^{m\times n}$ such that (6.15) holds if and only if there exist an $n \times n$ matrix $X > 0$ and a matrix $K \in \mathbb{R}^{m\times n}$ such that

$$0 > (A + BK)X + X(A + BK)^{\mathrm{T}}. \tag{6.16}$$

Now, defining $Z \triangleq KX$, it follows that there exist an $n \times n$ matrix $X > 0$ and a matrix $Z \in \mathbb{R}^{m \times n}$ such that

$$0 > AX + XA^{\mathrm{T}} + BZ + Z^{\mathrm{T}}B^{\mathrm{T}} \tag{6.17}$$

if and only if there exist an $n \times n$ matrix $X > 0$ and a matrix $K \in \mathbb{R}^{m \times n}$ such that (6.16) holds. Hence, the stabilizabiliy condition (6.15) holds if and only if the LMI condition (6.17) holds.

Definition 6.5 [9]. Consider the linear dynamical system given by (6.4). Then the pair (A, B) is *stabilizable-nonnegative orthant feedback holdable* if there exists a feedback control law of the form $u(t) = Kx(t)$ such that the closed-loop system is asymptotically stable and, for every initial condition $x_0 \in \overline{\mathbb{R}}_+^n$, $x(t) \in \overline{\mathbb{R}}_+^n$ for all $t \geq 0$.

Proposition 6.3. Consider the linear dynamical system (6.4). Then the following statements are equivalent:

i) The pair (A, B) is stabilizable-nonnegative orthant feedback holdable.

ii) There exists $K \in \mathbb{R}^{m \times n}$ such that $-(A + BK)$ is a nonsingular M-Matrix.

iii) There exist a diagonal $n \times n$ matrix $Y > 0$ and a matrix $K \in \mathbb{R}^{m \times n}$ such that $A + BK$ is essentially nonnegative and $(A + BK)^{\mathrm{T}}Y + Y(A + BK) < 0$.

Proof. The equivalence of statements *i*) and *ii*) follows from Theorem 4.1 of [9, p. 133]. Next, note that statement *i*) holds if and only if there exists $K \in \mathbb{R}^{m \times n}$ such that $A + BK$ is Hurwitz and essentially nonnegative or, equivalently, if and only if there exist a positive diagonal matrix $Y \in \mathbb{R}^{n \times n}$ and a matrix $K \in \mathbb{R}^{m \times n}$ such that $A + BK$ is essentially nonnegative and (6.15) holds. This proves the equivalence of statements *i*) and *iii*), and hence, statements *ii*) and *iii*). □

Next, we present an optimal control framework for minimizing an $\mathcal{H}_2$ norm bound of a nonnegative dynamical system while constraining the system states to the nonnegative orthant of the state space.

Problem 6.2(Static Compensation for Nonnegative Systems). Consider the linear dynamical system (6.9) with performance variables (6.11) where (A, B) is stabilizable-nonnegative orthant feedback holdable, A is essentially nonnegative, and B and x_0 are nonnegative. Determine K such that $A+BK$ is essentially nonnegative and Hurwitz, and the feedback control law $u(t) = Kx(t)$ minimizes the quadratic performance criterion (6.10).

The $\mathcal{H}_2$ optimal regulator problem for nonnegative dynamical systems given in Problem 6.2 is computationally intractable as an LMI. Specifically, if the solution to the GEVP (6.12)–(6.14) is such that $A + BK$ (with $K = ZX^{-1}$) is essentially nonnegative, then such a solution solves Problem 6.2. However, in general, $A+BK$ is not guaranteed to be essentially nonnegative for Problem 6.1, and the optimization problem (6.12)–(6.14) does not remain an LMI if the essential nonnegativity constraint on $A + BK$ is enforced. As noted above, the stabilizability condition (6.15) may be checked using the LMI condition (6.17). If we further require the essential nonnegativity of $A + BK$ or, equivalently, $(AX + BZ)X^{-1}$, the stabilizability condition will no longer remain an LMI. However, if X is positive and diagonal, then $AX + BZ$ is essentially nonnegative if and only if $(AX + BZ)X^{-1} = A + BK$ is essentially nonnegative. It follows from Proposition 6.3 that there exist a positive diagonal matrix $X \in \mathbb{R}^{n\times n}$ and a matrix $Z \in \mathbb{R}^{m\times n}$ such that (6.17) holds and $AX + BZ$ is essentially nonnegative, that is,

$$(AX + BZ)_{(i,j)} \geq 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, n,$$

if and only if (A, B) is stabilizable-nonnegative orthant feedback holdable. Based on this observation we propose the following $\mathcal{H}_2$ suboptimal control design framework for Problem 6.2.

Theorem 6.5. Consider the Static Compensation for Nonnegative Systems Problem given by Problem 6.2. Assume there exist matrices $X \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times n}$, $W \in \mathbb{R}^{p\times p}$, where X is diagonal and $W = W^{\mathrm{T}}$, such that (6.13) and (6.14) hold, and

$$0 \leq (AX + BZ)_{(i,j)}, \quad i \neq j, \quad i, j = 1, 2, \ldots, n. \tag{6.18}$$

Then $K = ZX^{-1}$ is such that $A + BK$ is Hurwitz and essentially nonnegative. Furthermore,

$$J(K_{\mathrm{opt}}) \leq J(K) < \operatorname{tr} W, \tag{6.19}$$

where K_{opt} denotes the solution to Problem 6.2. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.19) is given by

$$J(K_{\mathrm{opt}}) \leq J(K^*) \leq \operatorname{tr} W^*, \tag{6.20}$$

where $K^* = Z^*X^{*-1}$ and X^*, Z^*, and W^* are the optimal solutions to the GEVP (6.12) subject to (6.13), (6.14), and (6.18).

Proof. Note that (6.13) can be equivalently written as

$$0 > (A + BK)X + X(A + BK)^{\mathrm{T}} + DD^{\mathrm{T}}. \tag{6.21}$$

Hence, since $X > 0$ it follows that $A + BK$ is Hurwitz. Furthermore, since X is diagonal it follows from (6.18) that $A + BK = (AX + BZ)X^{-1}$ is essentially nonnegative. Next, using Schur complements, (6.14) holds if and only if $X > 0$ and

$$W > (E_1X + E_2Z)X^{-1}(E_1X + E_2Z)^{\mathrm{T}} = (E_1 + E_2K)X(E_1 + E_2K)^{\mathrm{T}}. \tag{6.22}$$

Now, with A replaced by $A + BK$ and E replaced by $E_1 + E_2K$ it follows from Theorem 6.3 that $\|\tilde{G}\|_2^2 < \operatorname{tr} W$, where $\tilde{G}(s) = (E_1 + E_2K)(sI_n - (A + BK))^{-1} D$ is the closed-loop system transfer function. Hence, $J(K) = \|\tilde{G}\|_2^2 < \operatorname{tr} W$. Next, (6.19) follows trivially since K is a feasible controller gain satisfying all of the constraints of Problem 6.2. Finally, (6.20) follows by noting that the minimal cost of (6.12) subject to (6.13), (6.14), and (6.18) with X diagonal will be higher than the minimal cost of (6.12) subject to (6.13), (6.14), and $A + BZX^{-1}$ essentially nonnegative without the restriction on X being diagonal. □

Remark 6.3. Problem 6.2 considered in Theorem 6.5 requires that the pair (A, B) is stabilizable-nonnegative orthant feedback holdable. Necessary and sufficient conditions for stabilizability and holdability in the nonnegative orthant via the feedback controller $u = Kx$ are given in Chapters 7 and 8 of [9].

Remark 6.4. Although stabilizable-nonnegative orthant feedback holdability is equivalent to an LMI condition with diagonal X, the computation of the $\mathcal{H}_2$ norm (and hence the optimal regulator control problem) of a nonnegative system cannot be formulated as an LMI.

Remark 6.5. Note that $J(\bar{K}) \leq J(K_{\text{opt}}) \leq J(K^*)$, where $\bar{K}$ and K_{opt} correspond to the solution to Problem 6.1 and Problem 6.2, respectively, and K^* is the optimal solution given in Theorem 6.5. This follows directly from Theorem 6.4 by noting that the minimal cost of (6.12) subject to (6.13), (6.14), and (6.18) will be higher than the minimal cost of (6.12) subject to (6.13) and (6.14) due to the additional constraint (6.18) and restriction on X being diagonal. Hence, Theorem 6.5 gives a framework for minimizing an upper bound on the $\mathcal{H}_2$ norm of the closed-loop system while preserving closed-loop system nonnegativity.

Remark 6.6. To develop static control laws for nonnegative systems with a nonnegative control input, it suffices to solve the optimization problem stated in Theorem 6.5 with the additional LMI constraint $Z \geq\geq 0$. To see this, note that $u(t) \geq\geq 0$, $t \geq 0$, if and only if $K \geq\geq 0$. Since $X > 0$ is diagonal, $K \geq\geq 0$ if and only if $Z = KX \geq\geq 0$.

Remark 6.7. The closed-loop dynamics for the system described in Problem 6.2 using the design method in Theorem 6.5 is nonnegative (whether or not A is essentially nonnegative and B is nonnegative) provided that $x(0) = x_0 \geq\geq 0$. In this case, it follows from Proposition 6.1 that $x(t) \geq\geq 0$, $t \geq 0$, with $w(t) \equiv 0$.

It is important to note that if $w(t)$ is a standard white noise disturbance signal, then $w(t)$ can take arbitrary large negative values, and hence, subvert the nonnegativity of the states of (6.9). To guarantee the nonnegativity of the state variables for the case where $w(t) \not\equiv 0$, the disturbance model $Dw(t)$ in (6.9) needs to be nonlinear with $w(\cdot) \in \mathcal{D} \subset \mathcal{L}_2$. To further elucidate this, recall that an equivalent characterization of the $\mathcal{H}_2$ norm of $G(s)$ is given by

$$|||G|||_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \|G(\jmath\omega)\|_{\mathrm{F}}^2 \mathrm{d}\omega,$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius matrix norm. In addition the following theorem is needed.

Theorem 6.6. Consider the linear dynamical system (6.5) and (6.6) where $w(\cdot) \in \mathcal{L}_2$. Then $|||z|||_\infty \leq |||G|||_2 |||w|||_2$, where $|||\cdot|||_\infty$ and $|||\cdot|||_2$ denote the $\mathcal{L}_\infty$ and $\mathcal{L}_2$ norms, respectively.

Proof. It follows from *ii*) of Corollary 3.1 of [33] that $|||z|||_\infty^2 \leq \sigma_{\max}(EQE^{\mathrm{T}})|||w|||_2^2$, where $Q \in \mathbb{R}^{n \times n}$ is the solution to the Lyapunov equation

$$0 = AQ + QA^{\mathrm{T}} + DD^{\mathrm{T}}.$$

Now, since $\sigma_{\max}(X) \leq \mathrm{tr}(X)$, $X \in \mathbb{R}^{p\times p}$, it follows that

$$|||z|||_\infty^2 \leq \sigma_{\max}(EQE^{\mathrm{T}})|||w|||_2^2 \leq \mathrm{tr}(EQE^{\mathrm{T}})|||w|||_2^2 = |||G|||_2^2|||w|||_2^2.$$

This completes the proof. □

To present an optimal control framework for minimizing an $\mathcal{H}_2$ norm *bound* of a nonnegative dynamical system while constraining the system states to the nonnegative orthant of the state space in the presence of system disturbances, we consider controlled dynamical systems of the form

$$\dot{x}(t) = Ax(t) + Bu(t) + DG(x(t))\hat{w}(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.23}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $D \in \mathbb{R}^{n\times d}$, $G : \mathbb{R}^n \to \mathbb{R}^{d\times q}$, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ is the control input, and $\hat{w}(t) \in \mathbb{R}^q$ is an $\mathcal{L}_2$ disturbance. Furthermore, we assume that $G(x)$ is such that $\sup_{x\in\mathbb{R}^n} \sigma_{\max}(G(x)) \leq 1$ and for every $i \in \{1, 2, \ldots, n\}$, if $x_i = 0$, then $\mathrm{row}_i(DG(x)) = 0$, where $\mathrm{row}_i(X)$ denotes the ith row of $X \in \mathbb{R}^{p\times q}$. Note that if $u = Kx$ is such that $A + BK$ is essentially nonnegative, then the solution to (6.23) is nonnegative for all $\hat{w}(\cdot) \in \mathcal{L}_2$ and $x_0 \in \overline{\mathbb{R}}_+^n$. In addition, note that if $\hat{w}(\cdot) \in \mathcal{L}_2$, then $w(t) \triangleq G(x)\hat{w}(t) \in \mathcal{L}_2$ for all $x \in \mathbb{R}^n$. Hence, (6.23) can be rewritten as (6.9) with $w(\cdot) \in \mathcal{D} \subset \mathcal{L}_2$, where

$$\mathcal{D} \triangleq \{w(\cdot) \in \mathcal{L}_2 \,:\, w(t) = G(x(t))\hat{w}(t),\ \hat{w}(\cdot) \in \mathcal{L}_2,$$
$$\text{and } x(t) \text{ is the solution to (6.23) for some } x_0 \in \overline{\mathbb{R}}_+^n\}.$$

Now, it follows from Theorem 6.6 that $|||\tilde{G}|||_2$ is an upper bound to the induced operator norm from $\mathcal{D}$ to $\mathcal{L}_\infty$, where $\tilde{G}(s)$ is the closed-loop transfer function from $\mathcal{L}_2$ disturbances $w(t)$ to $\mathcal{L}_\infty$ performance variables $z(t)$. In this case, Theorem 6.5 solves Problem 6.2 for the case where $w(\cdot) \in \mathcal{D} \subset \mathcal{L}_2$ in (6.9).

6.4. Suboptimal Estimation for Nonnegative Dynamical Systems

In this section, we present a suboptimal estimation framework for minimizing the $\mathcal{H}_2$ norm of the error dynamics of a linear dynamical system while constraining the estimated states to the nonnegative orthant of the state space. Note that if the nonnegativity constraint on the estimated states is relaxed, the $\mathcal{H}_2$ optimal solution is given by the least squares Kalman filter problem. Before stating the main theorem of this section, we state the standard least squares estimation problem and present a solution to this problem using LMIs.

Problem 6.3 ($\mathcal{H}_2$ Optimal Estimation Problem). Consider the linear dynamical system given by

$$\dot{x}(t) = Ax(t) + D_1 w(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.24}$$

$$y(t) = Cx(t) + D_2 w(t), \tag{6.25}$$

where $A \in \mathbb{R}^{n\times n}$, $D_1 \in \mathbb{R}^{n\times d}$, $D_2 \in \mathbb{R}^{l\times d}$, $C \in \mathbb{R}^{l\times n}$, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^l$ is the output measurement, and $w(t) \in \mathbb{R}^d$ is a standard white noise process. Furthermore, assume that (A, C) is detectable and $D_1 D_2^{\mathrm{T}} = 0$. Design an estimator of the form

$$\dot{x}_{\mathrm{e}}(t) = A_{\mathrm{e}} x_{\mathrm{e}}(t) + B_{\mathrm{e}} y(t), \quad x_{\mathrm{e}}(0) = x_{\mathrm{e}0}, \quad t \geq 0, \tag{6.26}$$

$$y_{\mathrm{e}}(t) = C_{\mathrm{e}} x_{\mathrm{e}}(t), \tag{6.27}$$

where $x_{\mathrm{e}}(t) \in \mathbb{R}^n$, $A_{\mathrm{e}} \in \mathbb{R}^{n\times n}$, $B_{\mathrm{e}} \in \mathbb{R}^{n\times l}$, and $C_{\mathrm{e}} \in \mathbb{R}^{n\times n}$, that satisfies the following design criteria: *i*) $x(t) - x_{\mathrm{e}}(t) \to 0$ as $t \to \infty$ when $w(t) \equiv 0$; and *ii*) the $\mathcal{H}_2$ error performance criterion

$$J(B_{\mathrm{e}}) = \lim_{t\to\infty} \mathbb{E}\left\{\frac{1}{t}\int_0^t z^{\mathrm{T}}(s) z(s) \mathrm{d}s\right\} \tag{6.28}$$

is minimized, where the weighted error $z(t)$ is given by $z(t) \stackrel{\triangle}{=} E\left(x(t) - x_{\mathrm{e}}(t)\right)$.

As in the standard Kalman filter design, we set $A_{\mathrm{e}} = A - B_{\mathrm{e}}C$ and $C_{\mathrm{e}} = I_n$ so that the only free parameter in Problem 6.3 is the Kalman gain B_{e}. Note that the constraint $A_{\mathrm{e}} = A - B_{\mathrm{e}}C$ automatically satisfies criterion $i)$ in Problem 6.3 by requiring that $A - B_{\mathrm{e}}C$ be Hurwitz. In this case, the error state $e(t) \to 0$ as $t \to \infty$ for $w(t) \equiv 0$, where $e(t) \triangleq x(t) - x_{\mathrm{e}}(t)$.

Theorem 6.7. Consider the $\mathcal{H}_2$ Optimal Estimation Problem given by Problem 6.3. Then the optimal Kalman gain and the optimal $\mathcal{H}_2$ cost are given by $B_{\mathrm{e}}^* = -Y^{*-1}V^{*\mathrm{T}}$ and $J(B_{\mathrm{e}}^*) = \operatorname{tr} U^*$, respectively, where $Y^* = Y^{*\mathrm{T}} \in \mathbb{R}^{n\times n}$, $U^* = U^{*\mathrm{T}} \in \mathbb{R}^{d\times d}$, and $V^* \in \mathbb{R}^{l\times n}$ are the optimal solutions to the GEVP

$$\inf_{Y\in\mathbb{R}^{n\times n},\, V\in\mathbb{R}^{l\times n},\, U\in\mathbb{R}^{d\times d}} \operatorname{tr} U \tag{6.29}$$

subject to

$$0 > A^{\mathrm{T}}Y + YA + C^{\mathrm{T}}V + V^{\mathrm{T}}C + E^{\mathrm{T}}E, \tag{6.30}$$

$$0 < \left[\begin{array}{cc} Y & (YD_1 + V^{\mathrm{T}}D_2) \\ (YD_1 + V^{\mathrm{T}}D_2)^{\mathrm{T}} & U \end{array}\right]. \tag{6.31}$$

Proof. The proof is dual to the proof of Theorem 6.4 and, hence, is omitted. □

Next, we present the optimal nonnegative estimator design problem for nonnegative systems.

Problem 6.4 (Optimal Estimation for Nonnegative Systems). Consider the linear dynamical system given by (6.24) and (6.25), and assume that $(A^{\mathrm{T}}, -C^{\mathrm{T}})$ is stabilizable-nonnegative orthant feedback holdable, A is essentially nonnegative, C and x_0 are nonnegative, and $D_1D_2^{\mathrm{T}} = 0$. Determine B_{e} such that $A - B_{\mathrm{e}}C$ is essentially nonnegative and Hurwitz, B_{e} is nonnegative, and the dynamic estimator (6.26) and (6.27) minimizes the least square error criterion (6.28).

Note that (6.24)–(6.26) can be rewritten as

$$\left[\begin{array}{c} \dot{x}(t) \\ \dot{x}_{\mathrm{e}}(t) \end{array}\right] = \left[\begin{array}{cc} A & 0 \\ B_{\mathrm{e}}C & A_{\mathrm{e}} \end{array}\right] \left[\begin{array}{c} x(t) \\ x_{\mathrm{e}}(t) \end{array}\right] + \left[\begin{array}{c} D_1 \\ B_{\mathrm{e}}D_2 \end{array}\right] w(t), \quad \left[\begin{array}{c} x(0) \\ x_{\mathrm{e}}(0) \end{array}\right] = \left[\begin{array}{c} x_0 \\ x_{\mathrm{e}0} \end{array}\right], \; t \geq 0. \tag{6.32}$$

Now, to guarantee that (6.32) is a nonnegative dynamical system with $w(t) \equiv 0$, we require that $A_{\mathrm{e}} = A - B_{\mathrm{e}}C$ is essentially nonnegative, $B_{\mathrm{e}} \geq\geq 0$, $x_0 \in \overline{\mathbb{R}}^n_+$, and $x_{\mathrm{e}0} \in \overline{\mathbb{R}}^n_+$.

Theorem 6.8. Consider the Optimal Estimation for Nonnegative Systems Problem given by Problem 6.4. Assume there exist matrices $Y \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{l\times n}$, and $U \in \mathbb{R}^{d\times d}$, where Y is diagonal and $U = U^{\mathrm{T}}$, such that (6.30) and (6.31) hold, and

$$\left(YA + V^{\mathrm{T}}C\right)_{(i,j)} \geq 0, \quad i \neq j, \quad i,j = 1,2,\ldots,n, \tag{6.33}$$

$$V \leq\leq 0. \tag{6.34}$$

Then $B_{\mathrm{e}} = -Y^{-1}V^{\mathrm{T}}$ is such that $A - B_{\mathrm{e}}C$ is Hurwitz and essentially nonnegative, and $B_{\mathrm{e}} \geq\geq 0$. Furthermore,

$$J(B_{\mathrm{e,opt}}) \leq J(B_{\mathrm{e}}) < \operatorname{tr} U, \tag{6.35}$$

where $B_{\mathrm{e,opt}}$ denotes the solution to Problem 6.4. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.35) is given by $J(B_{\mathrm{e,opt}}) \leq J(B^*_{\mathrm{e}}) \leq \operatorname{tr} U^*$, where $B^*_{\mathrm{e}} = -Y^{*-1}V^{*\mathrm{T}}$ and Y^*, U, and V^* are the optimal solutions to the GEVP (6.29) subject to (6.30), (6.31), (6.33), and (6.34).

Proof. Note that (6.30) can be equivalently written as

$$0 > (A - B_{\mathrm{e}}C)^{\mathrm{T}}Y + Y(A - B_{\mathrm{e}}C) + E^{\mathrm{T}}E. \tag{6.36}$$

Hence, since $Y > 0$ it follows that $A - B_{\mathrm{e}}C$ is Hurwitz. Furthermore, since Y is diagonal it follows from (6.33) that $A - B_{\mathrm{e}}C = Y^{-1}(YA + V^{\mathrm{T}}C)$ is essentially nonnegative. Next, $B_{\mathrm{e}} = -Y^{-1}V^{\mathrm{T}}$ is nonnegative if and only if (6.34) holds. The remainder of the proof now follows as in the proof of Theorem 6.5. $\square$

6.5. $\mathcal{H}_2$ Suboptimal Dynamic Controller Design for Nonnegative Dynamical Systems

In this section, we present a suboptimal dynamic controller design framework for minimizing the $\mathcal{H}_2$ norm of a linear dynamical system while constraining the plant states to the nonnegative orthant of the state space. The standard $\mathcal{H}_2$ dynamic compensation problem is given by the following problem.

Problem 6.5 ($\mathcal{H}_2$ Optimal Dynamic Compensation). Consider the linear dynamical system given by

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \quad x(0) = x_0, \quad t \geq 0, \tag{6.37}$$

$$y(t) = Cx(t) + D_2 w(t), \tag{6.38}$$

with performance variables (6.11), where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{l\times n}$, $D_1 \in \mathbb{R}^{n\times d}$, $D_2 \in \mathbb{R}^{l\times d}$, $E_1 \in \mathbb{R}^{p\times n}$, $E_2 \in \mathbb{R}^{p\times m}$, $x(t) \in \mathbb{R}^n$ is the system state vector, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^l$ is the output measurement, and $w(t) \in \mathbb{R}^d$ is a standard white noise process. Furthermore, assume that (A, B) is stabilizable, (A, E_1) is detectable, $E_1^{\mathrm{T}} E_2 = 0$, and $D_1 D_2^{\mathrm{T}} = 0$. Design a full-order dynamic compensator of the form

$$\dot{x}_{\mathrm{c}}(t) = A_{\mathrm{c}} x_{\mathrm{c}}(t) + B_{\mathrm{c}} y(t), \quad x_{\mathrm{c}}(0) = x_{\mathrm{c}0}, \quad t \geq 0, \tag{6.39}$$

$$u(t) = C_{\mathrm{c}} x_{\mathrm{c}}(t), \tag{6.40}$$

where $x_{\mathrm{c}}(t) \in \mathbb{R}^n$, $A_{\mathrm{c}} \in \mathbb{R}^{n\times n}$, $B_{\mathrm{c}} \in \mathbb{R}^{n\times m}$, and $C_{\mathrm{c}} \in \mathbb{R}^{l\times n}$, that satisfies the following design criteria: *i*) the undisturbed (i.e., $w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is asymptotically stable; and *ii*) the $\mathcal{H}_2$ performance criterion

$$J(A_{\mathrm{c}}, B_{\mathrm{c}}, C_{\mathrm{c}}) = \lim_{t\to\infty} \mathbb{E}\left\{ \frac{1}{t} \int_0^t z^{\mathrm{T}}(s) z(s) \mathrm{d}s \right\} \tag{6.41}$$

is minimized.

Theorem 6.9. Consider the $\mathcal{H}_2$ Optimal Dynamic Compensation Problem given by Problem 6.5. Then the optimal compensator gains are given by $A_{\mathrm{c}}^* = A + BC_{\mathrm{c}}^* - B_{\mathrm{c}}^* C$, $B_{\mathrm{c}}^* = -Y^{*-1}V^{*\mathrm{T}}$, and $C_{\mathrm{c}}^* = Z^* X^{*-1}$, where $X^* = X^{*\mathrm{T}} \in \mathbb{R}^{n\times n}$, $W^* = W^{*\mathrm{T}} \in \mathbb{R}^{p\times p}$, $Y^* = Y^{*\mathrm{T}} \in \mathbb{R}^{n\times n}$, $U^* = U^{*\mathrm{T}} \in \mathbb{R}^{d\times d}$, $V^* \in \mathbb{R}^{l\times n}$, and $Z^* \in \mathbb{R}^{m\times n}$ are the optimal solutions to the GEVP's (6.12) and (6.29) subject to (6.13), (6.14), and (6.30), (6.31), respectively.

Proof. The proof is a direct consequence of Theorems 6.4 and 6.7 by reducing the $\mathcal{H}_2$ Optimal Dynamic Compensation Problem to a combination of a full information problem and an output estimation problem. □

Next, we present the optimal dynamic compensation problem for nonnegative systems.

Problem 6.6 (Dynamic Compensation with Sign-Indefinite Input). Consider the linear dynamical system given by (6.37) and (6.38) where (A, B) and $(A^{\mathrm{T}}, -C^{\mathrm{T}})$ are sta- bilizable-nonnegative orthant feedback holdable, A is essentially nonnegative, B, C, and x_0 are nonnegative, $D_1 D_2^{\mathrm{T}} = 0$, and $E_1^{\mathrm{T}} E_2 = 0$. Determine the controller gains $(A_{\mathrm{c}}, B_{\mathrm{c}}, C_{\mathrm{c}})$ of (6.39) and (6.40) such that the undisturbed (i.e., $w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is asymptotically stable, the plant state $x(t) \in \overline{\mathbb{R}}_+^n$, $t \geq 0$, for $w(t) \equiv 0$, and the quadratic performance criterion (6.41) is minimized.

Theorem 6.10. Consider the Dynamic Compensation with Sign-Indefinite Input Problem given by Problem 6.6. Let $x_{\mathrm{c}0} = 0$ and assume there exist matrices $X \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times n}$, $W \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{l\times n}$, and $U \in \mathbb{R}^{d\times d}$, where X and Y are diagonal and $W = W^{\mathrm{T}}$ and $U = U^{\mathrm{T}}$, such that (6.13), (6.14), (6.30), (6.31) hold,

and

$$(AX + BZ)_{(i,j)} \geq 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, n, \tag{6.42}$$

$$BZ \leq\leq 0, \tag{6.43}$$

$$\big(YA + V^{\mathrm{T}}C\big)_{(i,j)} \geq 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, n. \tag{6.44}$$

Then $A_{\mathrm{c}} = A + BC_{\mathrm{c}} - B_{\mathrm{c}}C$, $B_{\mathrm{c}} = -Y^{-1}V^{\mathrm{T}}$, and $C_{\mathrm{c}} = ZX^{-1}$ are such that the undisturbed (i.e., $w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is asymptotically stable and plant state $x(t) \in \overline{\mathbb{R}}^n_+$ for all $x_0 \in \overline{\mathbb{R}}^n_+$ and $t \geq 0$. Furthermore,

$$J(A_{\mathrm{c,opt}}, B_{\mathrm{c,opt}}, C_{\mathrm{c,opt}}) \leq J(A_{\mathrm{c}}, B_{\mathrm{c}}, C_{\mathrm{c}}), \tag{6.45}$$

where $A_{\mathrm{c,opt}}$, $B_{\mathrm{c,opt}}$, and $C_{\mathrm{c,opt}}$ denote the solution to Problem 6.6. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.45) is given by

$$J(A_{\mathrm{c,opt}}, B_{\mathrm{c,opt}}, C_{\mathrm{c,opt}}) \leq J(A^*_{\mathrm{c}}, B^*_{\mathrm{c}}, C^*_{\mathrm{c}}), \tag{6.46}$$

where $A^*_{\mathrm{c}} = A + BC^*_{\mathrm{c}} - B^*_{\mathrm{c}}C$, $B^*_{\mathrm{c}} = -Y^{*-1}V^{*\mathrm{T}}$, and $C^*_{\mathrm{c}} = Z^*X^{*-1}$, and X^*, Z^*, Y^*, V^*, W^*, and U^* are the optimal solutions to the GEVP's (6.12) and (6.29) subject to (6.13), (6.14), (6.42), and (6.43), and (6.30), (6.31), and (6.44), respectively.

Proof. Note that the undisturbed ($w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is given by

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t), \quad \tilde{x}(0) = \tilde{x}_0, \quad t \geq 0, \tag{6.47}$$

where

$$\tilde{x} \triangleq \left[\begin{array}{c} x \\ x_{\mathrm{c}} \end{array} \right], \quad \tilde{x}_0 \triangleq \left[\begin{array}{c} x_0 \\ x_{\mathrm{c0}} \end{array} \right], \quad \tilde{A} \triangleq \left[\begin{array}{cc} A & BC_{\mathrm{c}} \\ B_{\mathrm{c}}C & A_{\mathrm{c}} \end{array} \right],$$

or, equivalently,

$$\left[\begin{array}{c} \dot{x}(t) \\ \dot{e}(t) \end{array} \right] = \left[\begin{array}{cc} A + BC_{\mathrm{c}} & -BC_{\mathrm{c}} \\ 0 & A - B_{\mathrm{c}}C \end{array} \right] \left[\begin{array}{c} x(t) \\ e(t) \end{array} \right], \quad \left[\begin{array}{c} x(0) \\ e(0) \end{array} \right] = \left[\begin{array}{c} x_0 \\ x_0 \end{array} \right], \; t \geq 0, \tag{6.48}$$

where $e(t) \triangleq x(t) - x_{\mathrm{c}}(t)$. Next, note that (6.13) and (6.30) can be equivalently written as

$$0 > (A + BC_{\mathrm{c}})X + X(A + BC_{\mathrm{c}})^{\mathrm{T}} + DD^{\mathrm{T}},$$
$$0 > (A - B_{\mathrm{c}}C)^{\mathrm{T}}Y + Y(A - B_{\mathrm{c}}C) + E^{\mathrm{T}}E.$$

Since $X > 0$ and $Y > 0$ it follows that $A + BC_{\mathrm{c}}$ and $A - B_{\mathrm{c}}C$ are Hurwitz, and hence, $\tilde{A}$ is Hurwitz. Furthermore, since X and Y are diagonal $A + BC_{\mathrm{c}} = (AX + BZ)X^{-1}$ and $A - B_{\mathrm{c}}C = Y^{-1}(YA + V^{\mathrm{T}}C)$ are essentially nonnegative. Moreover, $BC_{\mathrm{c}} \leq\leq 0$ holds if and only if (6.43) holds. Hence, it follows from (6.48) that $x(t) \in \overline{\mathbb{R}}^n_+$ for all $x_0 \in \overline{\mathbb{R}}^n_+$ and $t \geq 0$. The remainder of the proof follows using similar arguments as in the proof of Theorem 6.5. □

Remark 6.8. Note that in Problem 6.6, we only require that the plant states be nonnegative. Specifically, in the proof of Theorem 6.10 we show that the state $x(t)$ and the error state $e(t) = x(t) - x_{\mathrm{c}}(t)$ are nonnegative for $w(t) \equiv 0$, and hence, the compensator state $x_{\mathrm{c}}(t) = x(t) - e(t)$ need not be nonnegative. However, if $x(t) \geq\geq 0$ and $e(t)$ is "small," then $x_{\mathrm{c}}(t)$ is nonnegative.

In certain biological applications, control (source) inputs are usually constrained to be nonnegative as are the system states. Hence, next we develop dynamic control laws for nonnegative systems with nonnegative control inputs.

Problem 6.7 (Dynamic Compensation with Nonnegative Input). Consider the linear dynamical system given by (6.37) and (6.38), where (A, B) and $(A^{\mathrm{T}}, -C^{\mathrm{T}})$ are stabilizable-nonnegative orthant feedback holdable, A is essentially nonnegative, B, C, and x_0 are nonnegative, $D_1 D_2^{\mathrm{T}} = 0$, and $E_1^{\mathrm{T}} E_2 = 0$. Determine the controller gains $(A_{\mathrm{c}}, B_{\mathrm{c}}, C_{\mathrm{c}})$ of (6.39) and (6.40) such that the undisturbed (i.e., $w(t) \equiv 0$)

closed-loop system (6.37)–(6.40) is asymptotically stable, $x(t) \in \overline{\mathbb{R}}_+^n$ and $u(t) \in \overline{\mathbb{R}}_+^m$ for all $t \geq 0$ and $w(t) \equiv 0$, and the quadratic performance criterion (6.41) is minimized.

To guarantee the nonnegativity of the control input $u(t)$, $t \geq 0$, it follows from (6.40) that $C_{\mathrm{c}} x_{\mathrm{c}}(t)$ should be nonnegative. Here, we consider two different cases, namely, *i*) C_{c} and $x_{\mathrm{c}0}$ are nonnegative, and *ii*) C_{c} and $x_{\mathrm{c}0}$ are non-positive.

Theorem 6.11. Consider the Dynamic Compensation with Nonnegative Input Problem given by Problem 6.7. Let $x_{\mathrm{c}0} \geq\geq 0$ and assume there exist matrices $X \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times n}$, $W \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{l\times n}$, $U \in \mathbb{R}^{d\times d}$, where X and Y are diagonal and $W = W^{\mathrm{T}}$ and $U = U^{\mathrm{T}}$, such that (6.13), (6.14), (6.30), (6.31) hold, and

$$BZ \geq\geq 0, \tag{6.49}$$

$$\left(YA + V^{\mathrm{T}}C\right)_{(i,j)} \geq 0, \quad i \neq j, \quad i,j = 1,2,\ldots,n, \tag{6.50}$$

$$V^{\mathrm{T}}C \leq\leq 0. \tag{6.51}$$

Then $A_{\mathrm{c}} = A + BC_{\mathrm{c}} - B_{\mathrm{c}}C$, $B_{\mathrm{c}} = -Y^{-1}V^{\mathrm{T}}$, and $C_{\mathrm{c}} = ZX^{-1}$ are such that the undisturbed (i.e., $w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is asymptotically stable, and $x(t) \in \overline{\mathbb{R}}_+^n$ and $u(t) \in \overline{\mathbb{R}}_+^m$ for all $x_0 \in \overline{\mathbb{R}}_+^n$ and $t \geq 0$. Furthermore,

$$J(A_{\mathrm{c,opt}}, B_{\mathrm{c,opt}}, C_{\mathrm{c,opt}}) \leq J(A_{\mathrm{c}}, B_{\mathrm{c}}, C_{\mathrm{c}}), \tag{6.52}$$

where $A_{\mathrm{c,opt}}$, $B_{\mathrm{c,opt}}$, and $C_{\mathrm{c,opt}}$ denote the solution to Problem 6.7. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.52) is given by

$$J(A_{\mathrm{c,opt}}, B_{\mathrm{c,opt}}, C_{\mathrm{c,opt}}) \leq J(A_{\mathrm{c}}^*, B_{\mathrm{c}}^*, C_{\mathrm{c}}^*), \tag{6.53}$$

where $A_{\mathrm{c}}^* = A + BC_{\mathrm{c}}^* - B_{\mathrm{c}}^* C$, $B_{\mathrm{c}}^* = -Y^{*-1}V^{*\mathrm{T}}$, and $C_{\mathrm{c}}^* = Z^* X^{*-1}$, and X^*, Z^*, Y^*, V^*, W^*, and U^* are the optimal solutions to the GEVP's (6.12) and (6.29) subject to (6.13), (6.14), and (6.49), and (6.30), (6.31), (6.50), (6.51), respectively.

Proof. As in the proof of Theorem 6.10, it follows from (6.13) and (6.30) that $A + BC_{\rm c}$ and $A - B_{\rm c}C$ are Hurwitz, and hence, $\tilde{A}$ is Hurwitz. Next, since $Y > 0$ is diagonal, $(A - B_{\rm c}C) = Y^{-1}(YA + V^{\rm T}C)$ is essentially nonnegative, and $BC_{\rm c}$ and $B_{\rm c}C$ are nonnegative if and only if (6.49) and (6.51) hold. Hence, $\tilde{A}$ is essentially nonnegative. Thus, $\tilde{x}(t) \in \overline{\mathbb{R}}_+^{2n}$, $t \geq 0$, and hence, $x(t)$ and $x_{\rm c}(t)$ are nonnegative for all $t \geq 0$ and $w(t) \equiv 0$. Furthermore, (6.49) implies that $u(t) \in \overline{\mathbb{R}}_+^m$ for all $t \geq 0$. The remainder of the proof now follows using similar arguments as in the proof of Theorem 6.5. □

Theorem 6.12. Consider the Optimal Dynamic Compensation with Nonnegative Input Problem given by Problem 6.7. Let $x_{\rm c0} \leq\leq 0$ and assume there exist matrices $X \in \mathbb{R}^{n\times n}$, $Z \in \mathbb{R}^{m\times n}$, $W \in \mathbb{R}^{p\times p}$, $Y \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{l\times n}$, $U \in \mathbb{R}^{d\times d}$, where X and Y are diagonal and $W = W^{\rm T}$ and $U = U^{\rm T}$, such that (6.13), (6.14), (6.30), (6.31) hold, and

$$(AX + BZ)_{(i,j)} \geq 0, \quad i \neq j, \quad i, j = 1, 2, \ldots, n, \tag{6.54}$$

$$BZ \leq\leq 0, \tag{6.55}$$

$$V^{\rm T}C \geq\geq 0. \tag{6.56}$$

Then $A_{\rm c} = A + BC_{\rm c} - B_{\rm c}C$, $B_{\rm c} = -Y^{-1}V^{\rm T}$, and $C_{\rm c} = ZX^{-1}$ are such that the undisturbed (i.e., $w(t) \equiv 0$) closed-loop system (6.37)–(6.40) is asymptotically stable, and $x(t) \in \overline{\mathbb{R}}_+^n$ and $u(t) \in \overline{\mathbb{R}}_+^m$ for all $x_0 \in \overline{\mathbb{R}}_+^n$ and $t \geq 0$. Furthermore,

$$J(A_{\rm c,opt}, B_{\rm c,opt}, C_{\rm c,opt}) \leq J(A_{\rm c}, B_{\rm c}, C_{\rm c}), \tag{6.57}$$

where $A_{\rm c,opt}$, $B_{\rm c,opt}$, $C_{\rm c,opt}$ denote the solution to Problem 6.7. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.57) is given by

$$J(A_{\rm c,opt}, B_{\rm c,opt}, C_{\rm c,opt}) \leq J(A_{\rm c}^*, B_{\rm c}^*, C_{\rm c}^*), \tag{6.58}$$

where $A_{\mathrm{c}}^* = A + BC_{\mathrm{c}}^* - B_{\mathrm{c}}^* C$, $B_{\mathrm{c}}^* = -Y^{*-1}V^{*\mathrm{T}}$, and $C_{\mathrm{c}}^* = Z^* X^{*-1}$, and X^*, Z^*, Y^*, V^*, W^*, and U^* are the optimal solutions to the GEVP's (6.12) and (6.29) subject to (6.13), (6.14), (6.54), and (6.55) and (6.30), (6.31), and (6.56).

Proof. The proof is identical to the proof of Theorem 6.11 and, hence, is omitted. □

6.6. Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Suboptimal Control for Nonnegative Dynamical Systems

In this section, we present a mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ suboptimal control framework for nonnegative dynamical systems. Specifically, we introduce a static $\mathcal{H}_2$ state feedback control problem for nonnegative dynamical systems with constrained $\mathcal{H}_\infty$ disturbance attenuation between the plant disturbances and performance variables. First, however, we consider the standard mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ control problem [11] and present the solution to this problem using LMIs.

Problem 6.8(Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Optimal Regulator Problem). Given the linear controlled system (6.9) with performance variables (6.11) where (A, B) is stabilizable, determine a static state feedback control law $u(t) = Kx(t)$ that satisfies the following design criteria: *i*) the closed-loop system matrix given by $A + BK$ is Hurwitz; *ii*) the $\mathcal{H}_\infty$ norm of the closed-loop system from $\mathcal{L}_2$ disturbances $w(t)$ to $\mathcal{L}_2$ performance variables $z(t)$ satisfies

$$|||\tilde{G}(s)|||_\infty \leq \gamma, \tag{6.59}$$

where $|||\tilde{G}(s)|||_\infty \triangleq \sup_{\omega\in\mathbb{R}} \sigma_{\max}[\tilde{G}(\jmath\omega)]$ and $\gamma > 0$; and *iii*) the $\mathcal{H}_2$ performance criterion (6.10) is minimized.

The key step in enforcing the disturbance attenuation constraint (6.59) is to replace the Lyapunov inequality (6.13) by a Riccati inequality which overbounds X given by (6.13). The justification for this technique is provided by the following result.

Lemma 6.1 [11]. Consider the linear dynamical system $\mathcal{G}$ given by (6.5) and (6.6). Assume there exists $\mathcal{Q} = \mathcal{Q}^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ satisfying $\mathcal{Q} > 0$ and

$$0 > \begin{bmatrix} A\mathcal{Q} + \mathcal{Q}A^{\mathrm{T}} + DD^{\mathrm{T}} & (E\mathcal{Q})^{\mathrm{T}} \\ E\mathcal{Q} & -\gamma^2 I_n \end{bmatrix}. \tag{6.60}$$

Then A is Hurwitz, $\|\!|G(s)|\!\|_\infty \leq \gamma$, and $Q \leq \mathcal{Q}$, where Q satisfies (6.7), $\gamma > 0$, and $G(s) = E(sI_n - A)^{-1}D$. Consequently, $\operatorname{tr} EQE^{\mathrm{T}} \leq \operatorname{tr} E\mathcal{Q}E^{\mathrm{T}}$.

Theorem 6.13. Consider the Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Optimal Regulator Problem given by Problem 6.8. Assume there exist matrices $\mathcal{X} \in \mathbb{R}^{n\times n}$, $\mathcal{Z} \in \mathbb{R}^{m\times n}$, $\mathcal{W} \in \mathbb{R}^{p\times p}$, where $\mathcal{X} = \mathcal{X}^{\mathrm{T}}$ and $\mathcal{W} = \mathcal{W}^{\mathrm{T}}$, such that

$$0 < \begin{bmatrix} \mathcal{X} & (E_1\mathcal{X} + E_2\mathcal{Z})^{\mathrm{T}} \\ (E_1\mathcal{X} + E_2\mathcal{Z}) & \mathcal{W} \end{bmatrix}, \tag{6.61}$$

$$0 > \begin{bmatrix} A\mathcal{X} + \mathcal{X}A^{\mathrm{T}} + B\mathcal{Z} + \mathcal{Z}^{\mathrm{T}}B^{\mathrm{T}} + DD^{\mathrm{T}} & (E_1\mathcal{X} + E_2\mathcal{Z})^{\mathrm{T}} \\ E_1\mathcal{X} + E_2\mathcal{Z} & -\gamma^2 I_n \end{bmatrix}. \tag{6.62}$$

Then $K = \mathcal{Z}\mathcal{X}^{-1}$ is such that $A+BK$ is Hurwitz and (6.59) is satisfied. Furthermore,

$$J(K) \leq \mathcal{J}(K) < \operatorname{tr} \mathcal{W}, \tag{6.63}$$

where $\mathcal{J}(K) \triangleq (E_1 + E_2K)\mathcal{X}(E_1 + E_2K)^{\mathrm{T}}$. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.63) is given by

$$J(K^*) \leq \mathcal{J}(K^*) \leq \operatorname{tr} \mathcal{W}^*, \tag{6.64}$$

where $K^* = \mathcal{Z}^*\mathcal{X}^{*-1}$ and $\mathcal{X}^*$, $\mathcal{Z}^*$, and $\mathcal{W}^*$ are the optimal solutions to the GEVP

$$\inf_{\mathcal{X}\in\mathbb{R}^{n\times n},\, \mathcal{Z}\in\mathbb{R}^{m\times n},\, \mathcal{W}\in\mathbb{R}^{p\times p}} \operatorname{tr} \mathcal{W} \tag{6.65}$$

subject to (6.61) and (6.62).

Proof. Using Schur complements, (6.61) holds if and only if $\mathcal{X} > 0$ and

$$\mathcal{W} > (E_1\mathcal{X} + E_2\mathcal{Z})\mathcal{X}^{-1}(E_1\mathcal{X} + E_2\mathcal{Z})^{\mathrm{T}} = (E_1 + E_2K)\mathcal{X}(E_1 + E_2K)^{\mathrm{T}}. \quad (6.66)$$

Furthermore, (6.62) can be equivalently written as

$$0 > \left[\begin{array}{cc} (A + BK)\mathcal{X} + \mathcal{X}(A + BK)^{\mathrm{T}} + DD^{\mathrm{T}} & \mathcal{X}(E_1 + E_2K)^{\mathrm{T}} \\ (E_1 + E_2K)\mathcal{X} & -\gamma^2 I_n \end{array} \right]. \quad (6.67)$$

Now, with A replaced by $A+BK$ and E replaced by E_1+E_2K it follows from Lemma 6.1 that $A+BK$ is Hurwitz, (6.59) holds, and $\mathrm{tr}\,(E_1+E_2K)X(E_1+E_2K)^{\mathrm{T}} \leq \mathcal{J}(K)$, where X satisfies (6.13). Moreover, it follows from (6.66) and Theorem 6.3 that

$$J(K) = |||\tilde{G}(s)|||_2^2 \leq \mathcal{J}(K) = \mathrm{tr}\,(E_1 + E_2K)\mathcal{X}(E_1 + E_2K)^{\mathrm{T}} < \mathrm{tr}\,\mathcal{W}, \quad (6.68)$$

which proves (6.63). Finally, (6.64) is immediate. □

Next, we present an optimal control framework for minimizing an $\mathcal{H}_2$ norm bound of a nonnegative dynamical system while constraining the $\mathcal{H}_\infty$ norm of the closed-loop system and guaranteeing that the closed-loop system states remain in the nonnegative orthant of the state space.

Problem 6.9(Mixed $\mathcal{H}_2/\mathcal{H}_\infty$Compensation for Nonnegative Systems). Consider the linear dynamical system (6.9) with performance variables (6.11) where (A, B) is stabilizable-nonnegative orthant feedback holdable, A is essentially nonnegative, and B and x_0 are nonnegative. Determine K such that $A + BK$ is essentially nonnegative and Hurwitz, (6.59) holds, and the feedback control law $u(t) = Kx(t)$ minimizes the quadratic performance criterion (6.10).

Theorem 6.14. Consider the Mixed $\mathcal{H}_2/\mathcal{H}_\infty$ Compensation for Nonnegative Systems Problem given by Problem 6.9. Assume there exist matrices $\mathcal{X} \in \mathbb{R}^{n\times n}$, $\mathcal{Z} \in \mathbb{R}^{m\times n}$, $\mathcal{W} \in \mathbb{R}^{p\times p}$, where $\mathcal{X}$ is diagonal and $\mathcal{W} = \mathcal{W}^{\mathrm{T}}$, such that (6.61) and

(6.62) hold, and

$$0 \leq (A\mathcal{X} + B\mathcal{Z})_{(i,j)}, \quad i \neq j, \quad i, j = 1, 2, \ldots, n. \tag{6.69}$$

Then $K = \mathcal{Z}\mathcal{X}^{-1}$ is such that $A + BK$ is Hurwitz and essentially nonnegative. Furthermore,

$$J(K_{\mathrm{opt}}) \leq J(K) \leq \mathcal{J}(K) < \operatorname{tr} \mathcal{W}, \tag{6.70}$$

where K_{opt} denotes the solution to Problem 6.9 and $\mathcal{J}(K) = (E_1 + E_2K)\mathcal{X}(E_1 + E_2K)^{\mathrm{T}}$. Finally, the *sharpest* $\mathcal{H}_2$ bound satisfying (6.70) is given by $J(K_{\mathrm{opt}}) \leq J(K^*) \leq \mathcal{J}(K^*) \leq \operatorname{tr} \mathcal{W}^*$, where $K^* = \mathcal{Z}^*\mathcal{X}^{*-1}$ and $\mathcal{X}^*$, $\mathcal{Z}^*$, and $\mathcal{W}^*$ are the optimal solutions to the GEVP (6.65), subject to (6.61), (6.62), and (6.69).

Proof. The proof follows from Theorem 6.13 using similar arguments as in the proof of Theorem 6.5 and, hence, is omitted. □

The suboptimal $\mathcal{H}_2$ estimation and dynamic compensation frameworks for nonnegative dynamical systems considered in Sections 6.4 and 6.5 can be extended to the mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ case using similar techniques as presented above and the results given in [11, 12].

6.7. Illustrative Numerical Examples

In this section, we present two numerical examples for $\mathcal{H}_2$ suboptimal estimation and control for nonnegative dynamical systems.

Example 6.1. Consider the nonnegative dynamical system (6.9) with

$$A = \begin{bmatrix} -10 & 1 & 5 \\ 2 & -1 & 1 \\ 10 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{6.71}$$

Here, we design a suboptimal $\mathcal{H}_2$ controller for Problem 6.2. For our numerical simulation we take $x(0) = [2, 1, 1]^{\mathrm{T}}$, $E_1 = [I_3\, 0_{3\times 1}]^{\mathrm{T}}$, and $E_2 = [0_{3\times 1}\, 1]$. The YALMIP [100] and SeDuMi [137] MATLAB toolboxes are used to solve the LMI optimization problem given in Theorem 6.5. The $\mathcal{H}_2$ suboptimal control gain is given by $K = [-7.1440, -1, -8.4763]$. Figures 6.1–6.4 show the controlled states and control input versus time for an LQR design and the controller design given by Theorem 6.5. Clearly, the controlled states $x_1(t)$ and $x_3(t)$ take on negative values in the LQR design whereas all states remain nonnegative for the controller design given by Theorem 6.5. The optimal $\mathcal{H}_2$ cost is $J(K_{\mathrm{opt}}) = 6.9897$, whereas the $\mathcal{H}_2$ cost for the controller given by Theorem 6.5 is $J(K^*) = 8.4763$.

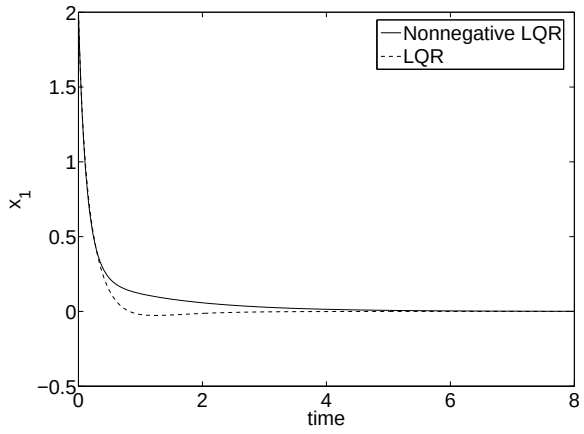

Figure 6.1: Comparison of $x_1(t)$ using the nonnegative LQR design and the LQR design

Example 6.2. In this example, we consider an estimation problem for the four compartment model for the disposition of the anesthetic drug propofol shown in Figure 6.5. This model was originally studied in [113]. A mass balance for the compartmental system yields a linear system of the form (6.24) and (6.25) (see [113]

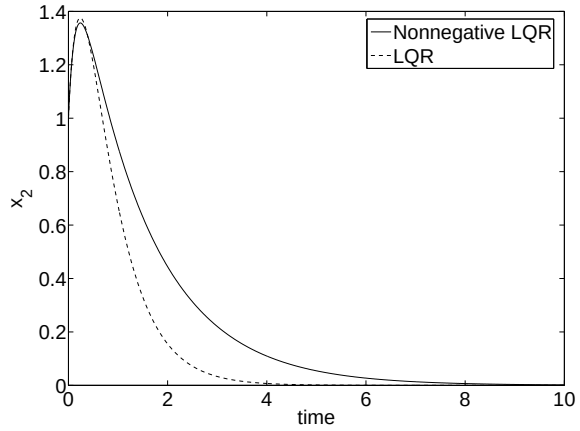

Figure 6.2: Comparison of $x_2(t)$ using the nonnegative LQR design and the LQR design

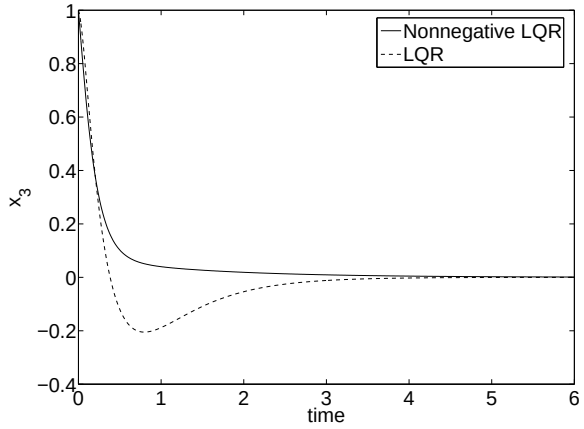

Figure 6.3: Comparison of $x_3(t)$ using the nonnegative LQR design and the LQR design

for the details) where

$$A = \begin{bmatrix} -3.4657 & 0.3114 & 0 & 0 \\ 0 & -0.3990 & 0.0920 & 0.0048 \\ 0 & 0.2070 & -0.0920 & 0 \\ 0 & 0.0400 & 0 & -0.0048 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^{\mathrm{T}}. \qquad (6.72)$$

Here, we assume that the system states as well as system measurements are driven by a standard white noise process so that $D_1 = \begin{bmatrix} I_4 & 0_{4\times 1} \end{bmatrix}$ and $D_2 = \begin{bmatrix} 0_{1\times 4} & 1 \end{bmatrix}$. For our numerical simulation we take $x(0) = [0.5, 0.5, 0.5, 0.5]^{\mathrm{T}}$, $x_{\mathrm{e}}(0) = [1, 1, 1, 1]^{\mathrm{T}}$, and a weighting matrix $E = I_4$. The YALMIP [100] and SeDuMi [137] MATLAB

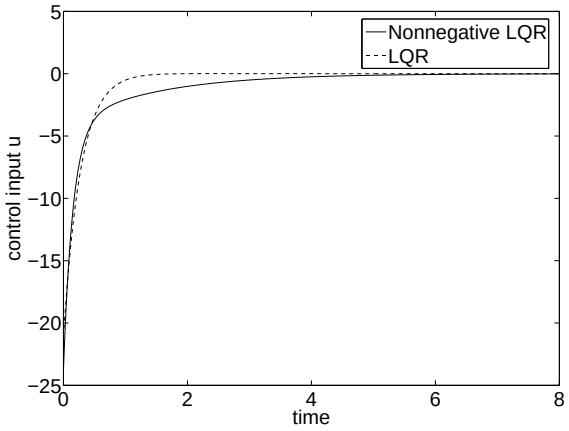

Figure 6.4: Comparison of $u(t)$ using the nonnegative LQR design and the LQR design

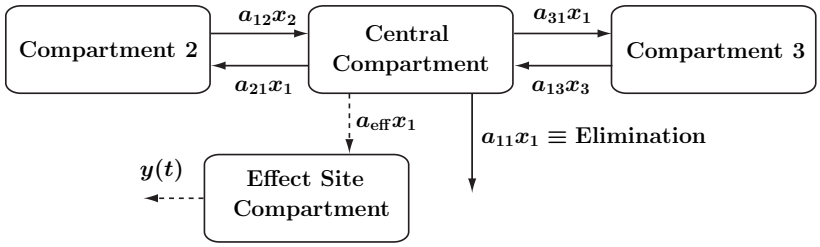

Figure 6.5: Four-compartment model for disposition of propofol

toolboxes are used to solve the LMI optimization problem given by Theorem 6.8. The estimator parameters are given by

$$A_{\text{e}} = \begin{bmatrix} -3.6974 & 0.3114 & 0 & 0 \\ 0 & -0.3990 & 0.0920 & 0.0048 \\ 0 & 0.2070 & -0.0920 & 0 \\ 0 & 0.0400 & 0 & -0.0048 \end{bmatrix}, \quad B_{\text{e}} = \begin{bmatrix} 0.2317 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (6.73)$$

Figures 6.6–6.9 show the actual states and estimated $\mathcal{H}_2$ optimal Kalman filter and $\mathcal{H}_2$ suboptimal nonnegative filter states of the system for the case where $w(t) \equiv 0$. Note that the estimated Kalman filter states take on negative values whereas all states remain nonnegative for the estimator design given by Theorem 6.8. The optimal $\mathcal{H}_2$ estimator error is $J(K_{\text{opt}}) = 137.01$, whereas the $\mathcal{H}_2$ estimator error for the estimator given by Theorem 6.8 is $J(K^*) = 563.1$.

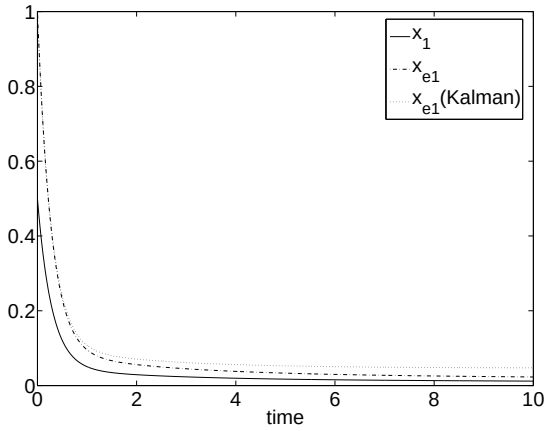

Figure 6.6: Comparison of $x_1(t)$ and $x_{\mathrm{e}1}(t)$ of the undisturbed system using standard and nonnegative Kalman filter design

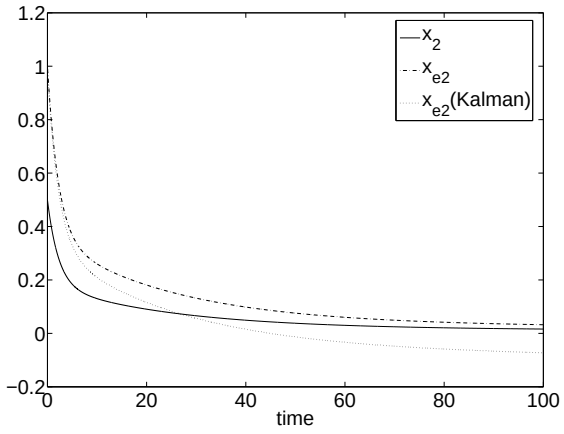

Figure 6.7: Comparison of $x_2(t)$ and $x_{\mathrm{e}2}(t)$ of the undisturbed system using standard and nonnegative Kalman filter design

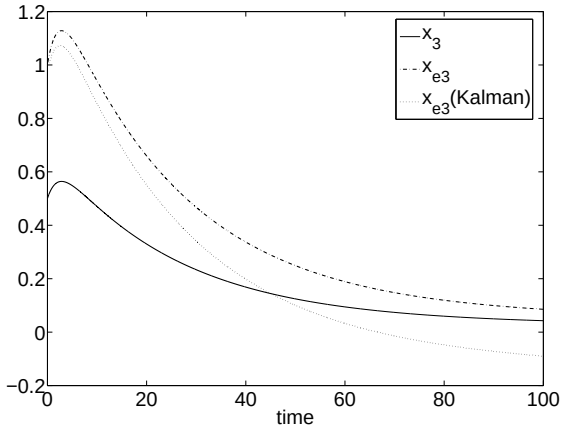

Figure 6.8: Comparison of $x_3(t)$ and $x_{e3}(t)$ of the undisturbed system using standard and nonnegative Kalman filter design

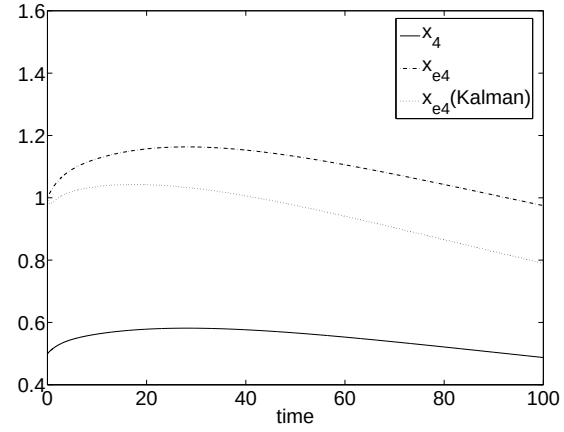

Figure 6.9: Comparison of $x_4(t)$ and $x_{e4}(t)$ of the undisturbed system using standard and nonnegative Kalman filter design

Chapter 7

Segmentation of the Endocardial Wall of the Left Atrium using Local Region-Based Active Contours and Statistical Shape Learning

7.1. Introduction

A powerful approach in medical image segmentation is active contour modeling wherein the boundaries of an object of interest are captured by minimizing an energy functional [16, 28]. The segmentation of the endocardial wall of the left atrium in delayed-enhancement magnetic resonance images (DE-MRI) using active contours is a challenging problem mainly due to the absence of clear boundaries. This usually leads either to contour "leaks," where the contour expands beyond the desired boundary, or partial segmentation, where the contour only captures the desired area partially. A shape-based segmentation approach can overcome this problem by using prior shape knowledge in the segmentation process. In this chapter, we use shape learning and shape-based image segmentation to identify the endocardial wall of the left atrium in the delayed-enhancement magnetic resonance images.

The outline of the chapter is as follows. In Section 7.2, we present the shape learning and shape-based image segmentation framework. Finally, in Section 7.3, this framework is applied to the problem of segmentation of the endocardial wall of the left atrium.

7.2. Shape Learning and Shape-Based Image Segmentation

In this section, we propose a shape-based image segmentation framework using the work of [141] and [97] to segment the endocardial wall of the left atrium. Our proposed approach involves two steps; namely, shape learning and image segmentation. To elucidate this, let the training set be composed of N binary images $B_i : \Omega \to \{0, 1\}$, $i = 1, \ldots, N$, where $\Omega \subset \mathbb{R}^3$ is the image domain. The binary image B_i corresponds to the segmentation of the endocardial wall of the left atrium for the image $I_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, N$, performed by a human expert. For $i = 1, \ldots, N$, $B_i(x) = 1$ if x falls inside the left atrial chamber and $B_i(x) = 0$ otherwise.

The first step in shape learning involves image registration. The goal in image registration is to align two given images, namely, the *fixed image* and the *moving image*, by finding a homeomorphism that maps the points in the moving image to the corresponding points in the fixed image. One of the most widely used techniques in image segmentation is the energy-based technique, where an energy functional describing the similarity between the two images is maximized (or, equivalently, an energy functional describing the discrepancy between the the two images is minimized) subject to a regularization constraint [140]. Here, we consider a special class of energy-based registration techniques, namely, the *mean-square-error affine registration scheme*, which is implemented in the insight segmentation and registration toolkit (ITK) [77]. While we consider binary images in this chapter, the mean-square-error affine registration scheme is applicable to any gray-scale image and is not limited to binary images.

Mean-Square-Error Affine Registration. Given the fixed image $B_{\mathrm{f}} : \Omega \to \{0, 1\}$ and the moving image $B_{\mathrm{m}} : \Omega \to \{0, 1\}$, where $\Omega \subset \mathbb{R}^3$, the goal of the mean-square-error affine registration scheme is to find an affine transformation $\mathcal{M} : \Omega \to \Omega$

such that the cost functional

$$J(A,T) \triangleq \int_{\Omega} \left(B_{\mathrm{m}}(\mathcal{M}(x)) - B_{\mathrm{f}}(x)\right)^2 \mathrm{d}x \tag{7.1}$$

is minimized, where $\mathcal{M}(x) \triangleq Ax + T$ with $A \in \mathcal{R}$ and $T \in \mathbb{R}^3$ denoting a rotation matrix and translation vector, respectively, and $\mathcal{R} \subset \mathbb{R}^{3\times 3}$ denoting the set of rotation matrices.

The mean-square-error affine registration scheme is an optimization problem which can be used to register the training images. Specifically, given the training set $\{B_1, \ldots, B_N\}$ with $B_i : \Omega \to \{0,1\}$, $i = 1, \ldots, N$, the mean-square-error affine registration scheme is used to register all binary images in the training set to an arbitrary binary image from the training set denoted by B_k, $k \in \{1, \ldots, N\}$. Hence, for $k \in \{1, \ldots, N\}$, B_k is regarded as the fixed image and B_i, $i = 1, \ldots, N$, $i \neq k$, is regarded as the moving image. This results in $N - 1$ mean-square-error registration problems given by

$$\min_{(A_i,T_i)\in\mathcal{R}\times\mathbb{R}^3} J_i^k(A_i,T_i), \quad A_i \in \mathbb{R}^{3\times 3},\ T_i \in \mathbb{R}^3,\ i = 1,\ldots,N,\ i \neq k, \tag{7.2}$$

where

$$J_i^k(A_i,T_i) \triangleq \int_{\Omega} \left(B_i(A_i x + T_i) - B_k(x)\right)^2 \mathrm{d}x.$$

We denote the registered binary moving images by $\hat{B}_i(x) \triangleq B_i(A_i^* x + T_i^*)$, $i = 1, \ldots, N$, $i \neq k$, and the binary fixed image by $\hat{B}_k(x) \triangleq B_k(x)$, $x \in \Omega$, where A_i^* and T_i^*, $i = 1, \ldots, N$, $i \neq k$, denote the optimal solutions to the optimization problem given by (7.2).

Next, we use principal component analysis (PCA) [88] to learn the registered shapes and create a statistical model for the shape. We need the following definition before stating the shape learning algorithm.

Definition 7.1. Given a closed surface $\mathcal{C}$ (which could correspond to the bound-

ary of a region of interest), the *signed distance function* $\phi : \Omega \to \mathbb{R}$ is the mapping defined by

$$\phi(x) \triangleq \begin{cases} \text{dist}(x, \mathcal{C}), & x \notin \mathcal{V}, \\ -\text{dist}(x, \mathcal{C}), & x \in \mathcal{V}, \end{cases} \tag{7.3}$$

where $\mathcal{V}$ is the volume enclosed by the closed surface $\mathcal{C}$ and $\text{dist}(\cdot, \cdot)$ is the distance operator defined by $\text{dist}(x, \mathcal{C}) \triangleq \inf_{y \in \mathcal{C}} \|x - y\|$, $x \in \Omega$.

In [141], the authors use the signed distance function to implicitly represent the training shapes, where the boundary of the training shape $\hat{B}_i(x)$, $x \in \Omega$, $i = 1, \ldots, N$, is given by the zero-level set of the signed distance function. In this chapter, we use a specific discrete approximation to the signed distance function, referred to as *sparse field level sets* (SFLS) [151], for the numerical implementation. Let $\Omega_{\text{s}} \subset \mathbb{R}^3$ denote the sampled image domain, where we use the $N_1 \times N_2 \times N_3$ image grid to sample Ω with N_j, $j = 1, 2, 3$, denoting the number of grid points in the j-th coordinate. In this case, the SFLS function $\Psi_i : \Omega_{\text{s}} \to \mathbb{R}$, $i = 1, \ldots, N$, associated with the binary image B_i satisfies

$$\Psi_i(x) > 0, \quad \text{if } x \notin \mathcal{A}_i, \tag{7.4}$$

$$\Psi_i(x) = 0, \quad \text{if } x \in \partial\mathcal{A}_i, \tag{7.5}$$

$$\Psi_i(x) < 0, \quad \text{if } x \in \mathcal{A}_i,\ x \notin \partial\mathcal{A}_i, \tag{7.6}$$

where $\mathcal{A}_i \triangleq \{x \in \Omega_{\text{s}} : B_i(x) = 1\}$ and $\partial\mathcal{A}_i$ denotes the boundary of $\mathcal{A}_i$. Sparse field level sets, which can be regarded as a variation of narrow-band methods [1], are an approximation to signed distance functions where the SFLS function assumes the same value as the signed distance function in the vicinity of the zero-level set. For points $x \notin \mathcal{A}_i$ (resp., $x \in \mathcal{A}_i$, $x \notin \partial\mathcal{A}_i$) which are sufficiently far from the zero-level set $\partial\mathcal{A}_i$, $\Psi_i(x) = 3$ (resp., $\Psi_i(x) = -3$).

Although the SFLS method was originally proposed in [151] to reduce the computational complexity of solving the partial differential equation governing the evolution

of the level set, we use SFLS for shape representation in the shape learning stage to control the variability of the level sets for points far from the zero-level set. Using SFLS to implicity represent the shape, as opposed to the more traditional signed distance function, is particularly important in PCA shape learning. Among all possible subspaces of a fixed dimension, PCA identifies the subspace in which the projection of the data has the maximum variance. Since we are only interested in the shape represented by the zero-level set, it is desired that PCA only reflect the variations in the zero-level set and not be influenced by the variations of the level set function at points far from the zero-level set. The SFLS function reduces this variability in the value of the level set functions $\Psi_1, \ldots, \Psi_N$ by assigning a constant value to points far from the level set. Note that SFLS representation is also used in the implementation of the localized region-based active contour described later in this section to reduce the computational complexity of the level set evolution.

We use the shape-learning framework proposed in [141] using PCA. First, the mean shape given by $\bar{\Phi} \triangleq \frac{1}{N}\sum_i^N \Psi_i$ is computed. Then, the mean-offset function is defined by $\tilde{\Psi}_i \triangleq \Psi_i - \bar{\Phi}$, $i = 1, \ldots, N$. Note that $\tilde{\Psi}_i : \Omega_{\mathrm{s}} \rightarrow \mathbb{R}$, $i = 1, \ldots, N$, where Ω_{s} is obtained by sampling Ω using the image grid. The $N_1 \times N_2 \times N_3$ image grid can be used to label each point in Ω_{s}, and hence, $\tilde{\Psi}_i$ can be transformed into an array of the size $N_1 \times N_2 \times N_3$. Next, we construct $s_i \in \mathbb{R}^M$, $i = 1, \ldots, N$, by forming a vector from the elements of the $N_1 \times N_2 \times N_3$ array associated with $\tilde{\Psi}_i$, where $M = N_1 \times N_2 \times N_3$ is the number of voxels in the binary image B_i, $i = 1 \ldots, N$. Define $S \triangleq [s_1, \ldots, s_N] \in \mathbb{R}^{M \times N}$ and $W \triangleq \frac{1}{N} S^{\mathrm{T}} S$. Finally, we use the Schur decomposition to obtain

$$W = Q^{\mathrm{T}} \Lambda Q, \tag{7.7}$$

where $Q \triangleq [q_1, \ldots, q_N]$, $q_i \in \mathbb{R}^N$, and $\Lambda \triangleq \mathrm{diag}[\lambda_1, \ldots, \lambda_N]$. The normalized eigenshapes Φ_i, $i = 1, \ldots, N$, are given by $\Phi_i = \frac{1}{\|Sq_i\|} Sq_i \in \mathbb{R}^M$, where we assume that

their corresponding eigenvalues λ_i, $i = 1, \ldots, N$, are in decreasing order, $Sq_i \neq 0$, $i = 1, \ldots, N$, and $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^M$.

In the framework proposed in [141], given a new image $I : \Omega \to \mathbb{R}$, it is assumed that the segmenting surface represented by a level-set function can be written as a weighted sum of the eigenshapes. Due to the complexity of the optimization problem and the presence of local minima, the initial guess in the optimization problem affects the optimal solution provided by the numerical algorithm. In this chapter, we introduce an intermediate step in which we use localized region-based active contours proposed in [97] to provide a better initialization to the optimization problem.

The localized region-based active contours provide a framework for segmenting heterogenous objects, where both global region-based and local edge-based methods fail [97]. The contour is implicitly represented by the signed distance function $\phi : \Omega \to \mathbb{R}$, where in our implementation we use the SFLS to represent ϕ due to its computational efficiency. Although we consider the Chan-Vese energy functional [31] within the localized region-based active contour framework, other energy functionals can be used [97].

For a given signed distance function ϕ define the *smoothed Heaviside function* $\mathcal{H}_\phi : \Omega \to \mathbb{R}$ by

$$\mathcal{H}_\phi(x) \triangleq \begin{cases} 1, & \phi(x) < -\epsilon, \\ 0, & \phi(x) > \epsilon, \\ \frac{1}{2}\left[1 + \frac{\phi}{\epsilon} + \frac{1}{\pi}\sin\left(\frac{\pi\phi(x)}{\epsilon}\right)\right], & \text{otherwise.} \end{cases} \tag{7.8}$$

In this case, the derivative $\delta_\phi : \Omega \to \mathbb{R}$ of the smoothed Heaviside function with respect to $x \in \Omega$ is given by

$$\delta_\phi(x) \triangleq \begin{cases} 0, & |\phi(x)| > \epsilon, \\ \frac{1}{2\epsilon}\left[1 + \cos\left(\frac{\pi\phi(x)}{\epsilon}\right)\right], & \text{otherwise.} \end{cases} \tag{7.9}$$

Moreover, for a given $r > 0$, define

$$\mathcal{B}_r(x, y) \triangleq \begin{cases} 1, & \|x - y\| < r, \\ 0, & \text{otherwise.} \end{cases} \tag{7.10}$$

The contour $\mathcal{C}$ segmenting the region of interest for a given image $I : \Omega \to \mathbb{R}$ is given by

$$\mathcal{C} = \{x \in \Omega : \phi^*(x) = 0\}, \tag{7.11}$$

where $\phi^* : \Omega \to \mathbb{R}$ is the solution of the minimization problem

$$\min_{\phi \in \mathbb{F}} E_{\text{lrac}}(\phi), \tag{7.12}$$

where $F \triangleq \{\phi : \Omega \to \mathbb{R} : \ \phi \ \text{is a signed distance function}\}$,

$$E_{\text{lrac}} \triangleq \int_\Omega \delta_\phi(x) \int_\Omega \mathcal{B}_r(x,y) F(I(y), \phi(y)) \mathrm{d}y \mathrm{d}x + \alpha \int_\Omega \delta_\phi(x) \|\nabla \phi(x)\| \mathrm{d}x, \quad \alpha > 0, \tag{7.13}$$

$$F(y) \triangleq \mathcal{H}_\phi(y)(I(y) - u_x)^2 + (1 - \mathcal{H}_\phi(y))(I(y) - v_x)^2, \tag{7.14}$$

$$u_x \triangleq \frac{\int_\Omega \mathcal{B}_r(x,y) \mathcal{H}_\phi(y) I(y) \mathrm{d}y}{\int_\Omega \mathcal{B}_r(x,y) \mathcal{H}_\phi(y) \mathrm{d}y}, \tag{7.15}$$

$$v_x \triangleq \frac{\int_\Omega \mathcal{B}_r(x,y)(1 - \mathcal{H}_\phi(y)) I(y) \mathrm{d}y}{\int_\Omega \mathcal{B}_r(x,y)(1 - \mathcal{H}_\phi(y) \mathrm{d}y}, \tag{7.16}$$

where $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^3$, ∇ denotes the gradient operator, and α is a regularization parameter. Here, we use the mean shape $\bar{\Phi}$ as the initialization for the localized region-based active contour.

Next, the result of the segmentation process is considered as an initial condition for the shape-based segmentation given in [141]. Specifically, define a new level set function $\Phi_{w, \mathcal{T}_p} : \Omega \to \mathbb{R}$ by

$$\Phi_{w, \mathcal{T}_p}(x) = \bar{\Psi}\left(\mathcal{T}_p(x)\right) + \sum_{i=1}^{e} \sqrt{\lambda_i} w_i \Phi_i\left(\mathcal{T}_p(x)\right), \tag{7.17}$$

where $w = [w_1 \ \ldots, \ w_e]^{\mathrm{T}}$, $w_i \in \mathbb{R}$, $i = 1, \ldots, e$, $e \in \mathbb{Z}_+$, $e < N$, is the number of selected eigenshapes, and $\mathcal{T}_p : \Omega \to \Omega$ is an affine transformation with parameter vector $p \in \mathbb{R}^7$ which includes translation, rotation, and magnification. Here we use the *binary mean model* [153] for the shape-based segmentation. Finally, note that the

segmented region is given by the zero-level set of the function $\Phi^*_{w,\mathcal{T}_p}$, where $\Phi^*_{w,\mathcal{T}_p}$ is the optimal solution of the optimization problem given by

$$\min_{(w,p)\in\mathbb{R}^k\times\mathbb{R}^7} E_{\text{binary}}(\Phi_{w,\mathcal{T}_p}), \tag{7.18}$$

where

$$E_{\text{binary}}(\Phi_{w,\mathcal{T}_p}) \triangleq -\frac{1}{2}\left(\frac{S_{\text{u}}}{A_{\text{u}}} - \frac{S_{\text{v}}}{A_{\text{v}}}\right)^2, \tag{7.19}$$

$$A_{\text{u}} \triangleq \int_\Omega \hat{\mathcal{H}}_{\Phi_{w,\mathcal{T}_p}}(y)\mathrm{d}y, \tag{7.20}$$

$$A_{\text{v}} \triangleq \int_\Omega (1-\hat{\mathcal{H}}_{\Phi_{w,\mathcal{T}_p}}(y))\mathrm{d}y, \tag{7.21}$$

$$S_{\text{u}} \triangleq \int_\Omega I(y)\hat{\mathcal{H}}_{\Phi_{w,\mathcal{T}_p}}(y)\mathrm{d}y, \tag{7.22}$$

$$S_{\text{v}} \triangleq \int_\Omega I(y)(1-\hat{\mathcal{H}}_{\Phi_{w,\mathcal{T}_p}}(y))\mathrm{d}y, \tag{7.23}$$

$$\hat{\mathcal{H}}_{\Phi_{w,\mathcal{T}_p}}(y) \triangleq \begin{cases} 1, & \Phi_{w,\mathcal{T}_p}(y) \leq 0, \\ 0, & \Phi_{w,\mathcal{T}_p}(y) > 0. \end{cases} \tag{7.24}$$

7.3. Application to Endocardial Wall Segmentation

In this section, we apply the framework of Section 7.2 to the problem of segmentation of the endocardial wall of the left atrium. Our data set includes 20 DE-MRI, namely $\{I_1,\ldots,I_{20}\}$, and their associated hand segmentations of the endocardial wall of the left atrium $B_1,\ldots,B_{20}$. These images are obtained from patients having undergone catheter ablation three months prior to the scan time. In our study, we use a hold-out method for cross-validation [15]. More specifically, the training set consists of the binary human-expert segmentations $B_6,\ldots,B_{20}$ and the test set consists of the DE-MR images $I_1,\ldots,I_5$. The segmentation results provided by the algorithm can be compared to the human-expert segmentations $B_1,\ldots,B_5$.

In the first step of the algorithm, the binary images in the training set $\{B_6,\ldots,B_{20}\}$ are registered to an arbitrary binary image, e.g., B_6. Next, the SFLS representation

described in the previous section is used to implicitly represent the training shapes. After subtracting the mean shape from all the training shapes, PCA learning is used to find the first 8 eigenshapes and their associated eigenvalues. This concludes the training phase of the algorithm.

Next, the trained algorithm is used for the segmentation of the endocardial wall of the left atrium by applying it to the test set $\{I_1, \ldots, I_5\}$. We assume that the human user can provide the algorithm with an approximate estimate of the centroid coordinates of the left atrium by a mouse click. Next, a translation transformation is applied to the mean shape so that its centroid coincides with the approximate centroid coordinates of the left atrium provided by the human user. The translated mean shape is used as an initialization for the localized region-based active contour algorithm. In the last step of the algorithm, the segmentation provided by the localized region-based active contour algorithm is used as an initialization for the shape-based segmentation algorithm described in the previous section. Table 7.1 outlines the proposed algorithm. The 3-dimensional and 2-dimensional view of the segmented endocardial wall of the left atrium for Patient 1 are given in Figure 7.1 and Figure 7.2, respectively.

Table 7.1: Endocardial Wall Segmentation Algorithm

Step 1.	Shape Learning.
a.	Register binary images $B_7, \ldots, B_{20}$ to B_6 using the mean-square-error affine registration scheme. Denote the registered training images by $\hat{B}_6, \ldots, \hat{B}_{20}$.
b.	Use SFLS to represent $\hat{B}_6, \ldots, \hat{B}_{20}$. Denote by $\Psi_6, \ldots, \Psi_{20}$.
c.	Compute the mean shape $\bar{\Phi}$. Compute the mean-offset functions $\tilde{\Psi}_6, \ldots, \tilde{\Psi}_{20}$.
d.	Construct $s_6, \ldots, s_{20}$ by forming a vector from the elements of $\tilde{\Psi}_6, \ldots, \tilde{\Psi}_{20}$.
e.	Compute $S = [s_1, \ldots, s_N]$ and $W = \frac{1}{N} S^{\mathrm{T}} S$.
f.	Use the Schur decomposition to obtain $W = Q^{\mathrm{T}} \Lambda Q$, where $Q = [q_1, \ldots, q_N]$ and $\Lambda = \mathrm{diag}[\lambda_1, \ldots, \lambda_N]$.
g.	Compute the normalized eigenshapes $\Phi_i = \frac{1}{\|Sq_i\|} Sq_i$.
h.	Select the first e eigenshapes corresponding to the largest eigenvalues. Denote by $\Phi_1, \ldots, \Phi_e$.
Step 2.	Image Segmentation.
a.	**for** $j = 1 : 5$
b.	Initialize the localized region-based active contour evolution for DE-MR image I_j using the mean-shape $\bar{\Phi}$ found in **Step 1c**.
c.	Evolve the segmenting surface $\mathcal{C}$ until some convergence criterion is met.
d.	Solve the optimization problem (7.18). Use the result of **Step 2c** as an initial guess for the optimization problem.
e.	**end for**

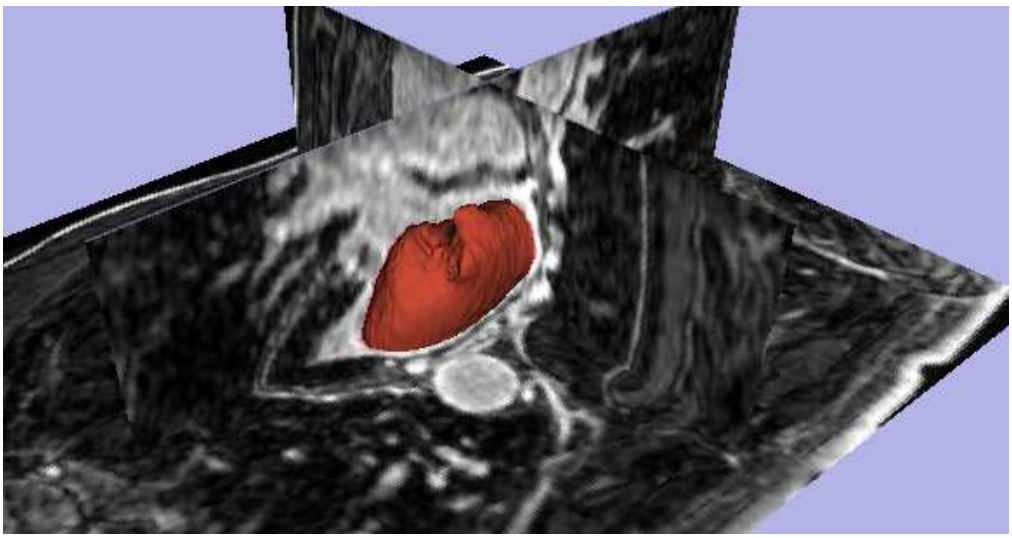

Figure 7.1: 3-dimensional view of the segmentation of the endocardial wall of the left atrium.

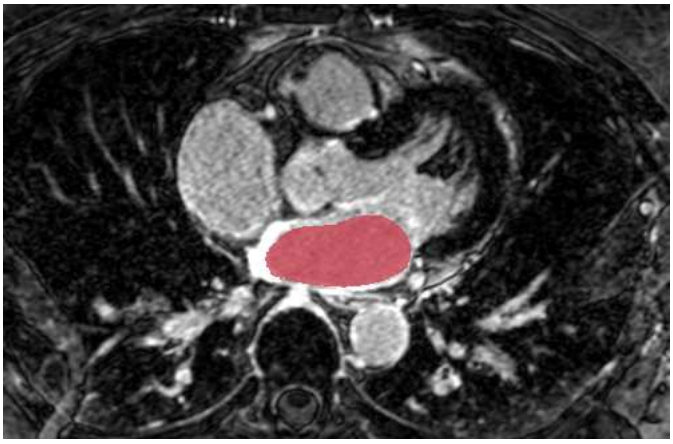

Figure 7.2: 2-dimensional view of the segmentation of the endocardial wall of the left atrium.

Chapter 8

Concluding Remarks and Recommendations for Future Research

8.1. Conclusion

In this dissertation, we presented an approach for designing clinical decision support and closed-loop control systems for cardiopulmonary management and sedation control in an intensive care unit using expert systems. It is important to note that expert systems are already in widespread use in other branches of medicine, more prominently in disease diagnosis, where the system inputs are the patient's details and symptoms, and the system outputs are probable diagnoses, recommended treatments or drugs which may be prescribed. Such systems are typically open-loop and may be regarded as rule-based search engines to help the clinician in his/her mapping of a given set of symptoms to a possible cause (disease).

Here, we are proposing to close the loop in a very specific sedation and cardiovascular function scenario using a set of heuristics in combination with Bayesian networks. A major challenge is the system identification aspect of the problem, that is, identifying a reasonable system model in case the plant is the patient. In contrast to more conventional identification techniques (e.g., sine sweeps), here the result must be more subjective but still very useful. Nevertheless, putting problems in drug administra-

tion in a closed-loop control framework has the strong potential of robustifying the therapies and making them much less amenable to human error.

Next, the problems of pain and pain intensity assessment using facial expressions in neonates were addressed. Sparse kernel machine algorithms were used to classify the images into pain and non-pain classes. The class membership posterior probability given by the relevance vector machine algorithm was interpreted as an estimate of the pain intensity, and this hypothesis was validated by comparing the results with expert and non-expert human assessments of pain. The results provided by the RVM algorithm can potentially be useful in decision support systems for ICU analgesia, where a reliable objective pain assessment measure is required.

We then considered facial expression recognition within an unsupervised learning framework. Specifically, given a data set composed of a number of facial images of the same subject with different facial expressions, the algorithm introduced in this dissertation segments the data set into groups corresponding to different facial expressions. Each facial image can be regarded as a point in a high-dimensional space, and the collection of images of the same subject resides on a manifold within this space. Our results show that different facial expressions reside on distinct subspaces if the manifold is unfolded. In particular, semi-definite embedding was used to reduce the dimensionality and unfold the manifold of facial images. Generalized principal component analysis was used to fit a series of subspaces to the data points and associate each data point to a subspace. Data points that belong to the same subspace were shown to belong to the same facial expression.

Next, we modeled the pharmacokinetics and pharmacodynamics of a general sedative agent using a hybrid deterministic-stochastic model involving deterministic pharmacokinetics and stochastic pharmacodynamics. Specifically, we used nonnegative and compartmental modeling to model the pharmacokinetics of propofol and a

stochastic process to represent the patient's sedation score and model the pharmacodynamics of propofol. Next, we used this deterministic-stochastic model to develop an open-loop optimal control policy for ICU sedation. Specifically, we first found the optimal effect-site drug concentration corresponding to a high probability for the desired sedation score (i.e., MRSS score of 3) and a low probability for all other sedation scores. Then, we used optimal control theory to drive the effect-site drug concentration to the optimal value found in the previous step while minimizing a given cost functional.

$\mathcal{H}_2$ suboptimal controllers using LMIs were developed for nonnegative dynamical systems. In particular, a series of generalized eigenvalue problems subject to a set of LMI constraints for designing $\mathcal{H}_2$ and mixed-norm $\mathcal{H}_2/\mathcal{H}_\infty$ suboptimal estimators, static controllers, and dynamic controllers were formulated for nonnegative dynamical systems. The proposed approach goes beyond nonnegative stabilization results in the literature by additionally addressing optimality considerations. Although the stability and optimality sufficient conditions presented in the dissertation are not necessary, our numerical simulations show that the performance of the suboptimal controller is close to that of the optimal controller while additionally ensuring nonnegativity of the plant states.

Finally, we proposed a shape-based image segmentation framework to segment the endocardial wall of the left atrium. The segmentation of the endocardial wall of the left atrium in delayed-enhancement magnetic resonance images using active contours is a challenging problem mainly due to the absence of clear boundaries. It was shown that a shape-based segmentation approach can overcome this problem by using prior shape knowledge in the segmentation process. Our proposed approach involved shape learning and image segmentation.

8.2. Recommendations for Future Research

The clinical decision support and closed-loop control framework presented in this dissertation will greatly benefit from advances in objective assessment of sedation and agitation. In future research, we propose to investigate the use of digital imaging and digital video of a patient's entire body movement as well as facial expressions to assess agitation and sedation in the ICU. In addition, using other sensor technologies such as an actiphraph and the bispectral index, as well as patient-ventilator interaction can allow for a more accurate and reliable assessment of patient's sedation and agitation condition. Correlations between the objective measurements for sedation and agitation using digital imaging and these other sensors, and a clinical standard assessment (e.g., MAAS or RASS score) needs to be investigated.

The open-loop optimal drug-dosing policy presented in this dissertation serves as a first step towards a general optimal closed-loop drug-dosing control policy to maintain a patient at a desired sedation state. We propose to extend the presented open-loop approach by closing the loop using the receding horizon control (RHC) [105]. In the RHC framework, closed-loop stability is acheived through the solution of a sequence of open-loop optimal control problems. Finding the corresponding sedative drug concentration in the effect-site compartment given the current measurement of the sedation score plays a key role in closing the loop. In addition, such an algorithm should guarantee stability given imperfect estimation of the drug concentration in the effect-site compartment. In addition, the drug dosage suggested by the algorithm should be compared to the drug dose prescribed by various experienced clinicians.

Markov chains are a powerful tool for modeling problems with a stochastic nature. Markov chains have been used in modeling, simulation, and also guideline and policy selection in a variety of disciplines, ranging from the value of maintenance within

a reliability theory framework [102] to medicine [90, 119]. The authors in [90, 119] have used Markov chains to model the severity of conditions of patients in ICU's and pediatric ICU's and used the model to calculate the expected duration of stay of patients.

Markov decision processes [76] are a theoretical framework for decision making in an uncertain and stochastic environment. In future research, we propose to model the sedation score evolution using Markov chains. The process of sedative drug concentration level selection will be modeled as a Markov decision process, where we can use drug concentration as a control (action) variable to regulate the sedation score around a desired predetermined score. Once such drug concentration levels are determined, the appropriate drug dosing policy (in terms of drug infusion rates) can be found using the theory of compartmental dynamical systems [61].

Finally, we propose to extend the approach presented in this dissertation to segment the epicardial wall, and ultimately, the automatic segmentation of the atrial wall. In addition, the enhanced regions in the DE-MRI can be assessed using computer aided statistical techniques. This can greatly benefit the study of ablation therapy for atrial fibrillation patients.

References

[1] D. Adalstein and J. A. Sethian, "A fast level set method for propogating interfaces," *J. Comput. Phys.*, vol. 118, pp. 269–277, 1995.

[2] J. M. Bailey, "Management of the hemodynamically unstable patient," in *Current Reviews in Clinical Anesthesia* (F. Moya, ed.), vol. 20, ch. 12, 2000.

[3] J. M. Bailey and W. M. Haddad, "Drug dosing control in clinical pharamacology: Paradigms, benefits, and challenges," *Contr. Sys. Mag.*, vol. 25, pp. 35–51, 2005.

[4] J. Barr, T. D. Egan, N. F. Sandoval, K. Zomorodi, C. Cohane, P. L. Gambus, and S. L. Shafer, "Propofol dosing regimens for ICU sedation based upon an integrated pharmacokinetic pharmacodynamic model," *Anesthesiology*, vol. 95, pp. 324–333, 2001.

[5] J. Barr, K. Zomorodi, E. J. Bertaccini, S. L. Shafer, and E. Geller, "A double–blind, randomized comparison of IV Lorazepam versus Midazolam for sedation of ICU patients via a pharmacologic model," *Anesthesiology*, vol. 95, pp. 286–298, 2001.

[6] P. Becouze, C. Hann, J. Chase, and G. Shaw, "Measuring facial grimacing for quantifying patient agitation in critical care," *Comp. Meth. Programs Biomed.*, vol. 87, pp. 138–147, 2007.

[7] R. Bellman, "Topics in pharmacokinetics. I. Concentration dependent rates," *MAB*, vol. 6, pp. 13–17, 1970.

[8] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer, 1985.

[9] A. Berman, M. Neumann, and R. J. Stern, *Nonnegative Matrices in Dynamic Systems*. New York: Wiley and Sons, 1989.

[10] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic Press, 1979.

[11] D. S. Bernstein and W. M. Haddad, "LQG control with an $\mathcal{H}_\infty$ performance bound: A Riccati equation approach," *IEEE Trans. Autom. Control*, vol. 34, pp. 293–305, 1989.

[12] D. S. Bernstein and W. M. Haddad, "Steady-state Kalman filtering with an $\mathcal{H}_\infty$ error bound," *Syst. Control Lett.*, vol. 12, pp. 9–16, 1989.

[13] D. S. Bernstein and D. C. Hyland, "Compartmental modeling and second-moment analysis of state space systems," *SMA*, vol. 14, pp. 880–901, 1993.

[14] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press, 1995.

[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2006.

[16] A. Blake and M. Isard, *Active Contours*. New York, NY: Springer-Verlag, 1998.

[17] B. Bonroy, G. Leysens, D. Miljkovic, P. Schiepers, E. Triau, M. Wils, D. Berckmans, P. Colleman, L. D. Maesschalck, S. Quanten, and B. Vanrumste, "Image acquisition system to monitor discomfort in demented elderly patients," in *Proc. Third Europ. Conf. Use of Modern Inform. Comm. Tech.*, Ghent, Belgium, 2008.

[18] S. P. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[20] S. Brahnam, C.-F. Chuang, S. Randall, and F. Shih, "Machine assessment of neonatal facial expressions of acute pain," *Dec. Supp. Sys.*, vol. 43, pp. 1242–1254, 2007.

[21] S. Brahnam, C.-F. Chuang, F. Shih, and M. Slack, "Machine recognition and representation of neonatal facial displays of acute pain," *Artif. Intel. Med.*, vol. 36, pp. 211–222, 2006.

[22] S. Brahnam, L. Nanni, and R. Sexton, "Introduction to neonatal facial pain detection using common and advanced face classification techniques," *Stud. Comput. Intel.*, vol. 48, pp. 225–253, 2007.

[23] S. Brahnam, L. Nanni, and R. Sexton, "Neonatal facial pain detection using NNSOA and LSVM," in *Proc. Int. Conf. Image Proc. Comput. Vis. Patt. Recog.*, Las Vegas, NV, 2008.

[24] M. A. Brodsky, B. J. Allen, C. J. Walker, T. P. Casey, C. R. Luckett, and W. L. Henry, "Amiodarone for maintenance of sinus rhythm after conversion of atrial fibrillation in the setting of a dilated left atrium," *Am. J. Cardiol.*, vol. 60, pp. 572–575, 1987.

[25] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. Waltham, MA: Ginn and Company, 1969.

[26] B. G. Buchanan and E. H. Shortliffe, *The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley, 1984.

[27] E. R. Carson, C. Cobelli, and L. Finkelstein, *The Mathematical Modeling of Metabolic and Endocrine Systems*. New York: Wiley, 1983.

[28] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comp. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[29] E. Castillo, J. M. Gutierrez, and A. S. Hadi, *Expert Systems and Probabilistic Network Models*. New York, NY: Springer, 1997.

[30] Y. Chang, C. Hu, R. Feris, and M. Turk, "Manifold based analysis of facial expression," *Imag. Vis. Comput.*, vol. 24, pp. 605–614, 2006.

[31] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Imag. Proces.*, vol. 10, pp. 266–277, 2001.

[32] J. G. Chase, F. Agogue, C. Starfinger, Z. Lam, G. M. Shaw, A. D. Rudge, and H. Sirisena, "Quantifying agitation in sedated ICU patients using digital imaging," *Comp. Meth. and Prog. in Biomed.*, vol. 76, no. 2, pp. 131–141, 2004.

[33] V. Chellaboina, W. M. Haddad, D. S. Bernstein, and D. A. Wilson, "Induced convolution operator norms of linear dynamical systems," *Math. Control Signals Syst.*, vol. 13, pp. 216–239, 2000.

[34] Y. Cherruault, *Mathematical Modelling in Biomedicine*. Reidel: Dordrecht, 1986.

[35] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, pp. 37–46, 1960.

[36] J. Cohen, "Weighted kappa: Nominal scale agreement with provision for scale disagreement or partial credit," *Psych. Bull.*, vol. 70, pp. 213–220, 1968.

[37] K. D. Craig, H. D. Hadjistavropoulos, R. V. E. Grunau, and M. F. Whitfield, "A comparison of two measures of facial activity during pain in the newborn child," *J. Pediat. Psych.*, vol. 19, pp. 305–318, 1994.

[38] K. D. Craig, K. M. Prkachin, and R. V. E. Grunau, "The facial expression of pain," in *Handbook of Pain Assessment* (D. Turk and R. Melzack, eds.), New York, NY: Guilford, 2001.

[39] H. J. Crijns, I. C. V. Gelder, H. J. V. der Woude, J. G. Grandjean, R. G. Tieleman, J. Brugemann, P. J. D. Kam, and T. Ebels, "Efficacy of serial electrical cardioversion therapy in patients with chronic atrial fibrillation after valve replacement and implications for surgery to cure atrial fibrillation," *Am. J. Cardiol.*, vol. 78, pp. 1140–1144, 1996.

[40] J. Devlin, G. Boleski, M. Mlynarek, D. Nerenz, E. Peterson, M. Jankowski, H. Horst, and B. Zarowitz, "Motor activity assessment scale: A valid and reliable sedation scale for use with mechanically ventilated patients in an adult surgical intensive care unit," *Crit. Care Med.*, vol. 1, pp. 1271–1275, 1999.

[41] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory: A Convex Approach*. New York: Springer-Verlag, 2000.

[42] Y. Ebihara and T. Hagiwara, "New dilated LMI characterizations forcontinuous-time multiobjective controller synthesis," *Automatica*, vol. 40, pp. 2003–2009, 2004.

[43] R. G. Eckenhoff and J. S. Johansson, "On the relevance of "clinically relevant concentrations" of inhaled anesthetics in in vitro experiments," *Anesthesiology*, vol. 91, no. 3, pp. 856–860, 1999.

[44] D. Eisenbud, *Commutative Algebra*. Springer-Verlag, 1995.

[45] P. Ekman, *Emotion in the Human Face*. New York, NY: Cambridge University Press, 1982.

[46] B. S. Everitt, *Statistical Methods in Medical Investigations*. London: Edward Arnold, 1994.

[47] R. H. Falk, "Atrial fibrillation," *N. Engl. J. Med.*, vol. 344, no. 14, pp. 1067–1078, 2001.

[48] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*. New York, NY: John Wiley & Sons, 2000.

[49] E. Feron and F. Alegre, "Control software analysis, part II: Closed-loop analysis," 2008, Available from `http://www.citebase.org/abstract?id=oai:arXiv.org:0812.1986`.

[50] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Mach. Learn.*, vol. 29, pp. 131–163, 1997.

[51] J. A. Gallian, *Contemporary Abstract Algebra*. D. C. Heath and Company, 1990.

[52] I. C. V. Gelder, H. J. Crijns, R. G. Tieleman, J. Brugemann, P. J. D. Kam, A. T. Gosselink, F. W. Verheugt, and K. I. Lie., "Chronic atrial fibrillation: Success of serial cardioversion therapy and safety of oral anticoagulation," *Arch. Intern. Med.*, vol. 156, no. 22, pp. 2585–2592, 1996.

[53] B. Gholami, W. M. Haddad, and A. Tannenbaum, "Agitation and pain assessment using digital imaging," in *Proc. IEEE Eng. Med. Biolog. Conf.*, Minneapolis, MN, pp. 2176–2179, 2009.

[54] B. Gholami, W. M. Haddad, and A. Tannenbaum, "Relevance vector machine learning for neonate pain intensity assessment using digital imaging," *IEEE Trans. Biomed. Eng.*, vol. 57, pp. 1457–1466, 2010.

[55] A. G. Gilman, J. G. Hardman, and L. E. Limbird, *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. New York, NY: McGraw-Hill, 10 ed., 1996.

[56] P. S. Glass, M. Bloom, L. Kearse, C. Rosow, P. Sebel, and P. Manberg, "Bispectral analysis measures sedation and memory effects of propofol, midazolam, isoflurane, and alfentanil in normal volunteers," *Anesthesiology*, vol. 86, no. 4, pp. 836–847, 1997.

[57] M. J. Grap, T. Borchers, C. L. Munro, R. K. Elswick, and C. N. Sessler, "Actigraphy in the critically ill: Correlation with acitivity, agition, and sedation," *Amer. J. Crit. Care*, vol. 14, no. 1, pp. 52–60, 2005.

[58] F. S. Grodins, *Control Theory and Biological Systems*. New York: Columbia University Press, 1963.

[59] P. V. Grootendorst, D. H. Feeny, and W. Furlong, "Does it matter whom and how you ask? Inter- and intra-rater agreement in Ontario health survey," *J. Clin. Epidemiol.*, vol. 50, pp. 127–135, 1997.

[60] W. M. Haddad and J. M. Bailey, "Closed-loop control for intensive care unit sedation," *Best Pract. Res. Clin. Anaesth.*, vol. 23, pp. 95–114, 2009.

[61] W. M. Haddad, V. Chellaboina, and Q. Hui, *Nonnegative and Compartmental Dynamical Systems*. Princeton, NJ: Princeton University Press, 2010.

[62] W. M. Haddad and V. Chellaboina, "Stability and dissipativity theory for nonnegative dynamical systems: A unified analysis framework for biological and physiological systems," *Math. Biosci.*, to appear.

[63] W. M. Haddad, T. Hayakawa, and J. M. Bailey, "Adaptive control for nonnegative and compartmental dynamical systems with applications to general anesthesia," *Int. J. Adapt. Control Signal Process.*, vol. 17, pp. 209–235, 2003.

[64] T. Hadjistavropoulos and K. D. Craig, "A theoretical framework for understanding self-report and observational measures of pain: A communications model," *Behav. Res. Ther.*, vol. 40, pp. 551–570, 2002.

[65] F. Harary, *Graph Theory*. Reading, MA: Addison-Wesley, 1969.

[66] D. W. Hawes, J. A. Ross, D. C. White, and R. T. Wloch, "Servocontrol of closed circuit anesthesia," *Brit. J. Anaesth.*, vol. 54, pp. 229–230, 1982.

[67] T. Hayakawa, W. M. Haddad, J. M. Bailey, and N. Hovakimyan, "Passivity-based neural network adaptive output feedback control for nonlinear nonnegative dynamical systems," *IEEE Trans. Neural Networks*, pp. 387–398, 2005.

[68] T. Hayakawa, W. M. Haddad, and J. M. Baily, "Adaptive control for nonlinear compartmental dynamical systems with applications to clinical pharmacology," *Syst. Control Lett.*, vol. 55, pp. 62–70, 2006.

[69] T. Hayakawa, W. M. Haddad, J. M. Baily, and N. Hovakimyan, "Passivity-based neural network adaptive output feedback control for nonlinear nonnegative dynamical systems," *IEEE Trans. Neural Networks*, vol. 16, pp. 387–398, 2005.

[70] T. Hayakawa, W. M. Haddad, N. Hovakimyan, and V. Chellaboina, "Neural network adaptive control for nonlinear nonnegative dynamical systems," *IEEE Trans. Neural Networks*, vol. 16, pp. 399–413, 2005.

[71] J. K. Hayes, D. R. Westenkow, T. D. East, and W. S. Jordan, "Computer controlled anesthesia delivery system," *Med. Instrument*, vol. 18, pp. 224–231, 1984.

[72] M. N. Hijazi and A. Chaudhary, "The correlation between the bispectral index (bis) and motor activity assessment scale (maas) in sedated and mechanically ventilated patients," *CHEST Meeting Abstracts*, p. 897s, 2004.

[73] A. V. Hill, "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves," *J. Physiol.*, vol. 40, no. 1, pp. iv–vii, 1910.

[74] A. A. Hopgood, *Intelligent Systems for Engineers and Scientists*. Boca Raton, FL: CRC Press, 2001.

[75] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith, "Effects of computer-based clinical decision support systems of physician performance and patient outcomes," *J. Am. Med. Assoc.*, vol. 280, pp. 1339–1346, 1998.

[76] Q. Hu and W. Yue, *Markov Decision Processes with Their Application*. New York, NY: Springer, 2008.

[77] L. Ibanez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide*. second ed., 2005. http://www.itk.org/ItkSoftwareGuide.pdf.

[78] M. Imhoff and S. Kuhls, "Alarm algorithms in critical care monitoring," *Anesth. Analg.*, vol. 102, pp. 1525–1537, 2006.

[79] E. F. Ismail, S. J. Kim, and M. R. Salem, "Direct effects of propofol on myocardial contractility in situ canine hearts," *Anesthesiology*, vol. 79, no. 5, pp. 964–972, 1992.

[80] P. Jackson, *Introduction to Expert Systems*. Harlow, England: Addison-Wesley, 1999.

[81] J. Jacobi, G. L. Fraser, D. B. Coursin, R. R. Riker, D. Fontaine, E. T. Wittbrodt, D. B. Chalfin, M. F. Masica, H. S. Bjerke, W. M. Coplin, D. W. Crippen, B. D. Fuchs, R. M. Kelleher, P. E. Marik, S. A. Nasraway, M. J. Murray, W. T. Peruzzi, and P. D. Lumb, "Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult," *Crit. Care Med.*, vol. 30, pp. 119–141, 2002.

[82] J. A. Jacquez, *Compartmental Analysis in Biology and Medicine*. Ann Arbor, MI: University of Michigan Press, 1985.

[83] J. A. Jacquez, *Compartmental Analysis in Biology and Medicine, 2nd ed.* Ann Arbor: University of Michigan Press, 1985.

[84] J. A. Jacquez, *Modeling with Compartments*. Ann Aarbor, MI: BioMedWare, 1999.

[85] J. A. Jacquez and C. P. Simon, "Qualitative theory of compartmental systems," *SIAM Rev.*, vol. 35, pp. 43–79, 1993.

[86] P. Jais, R. Weerasooriya, D. C. Shah, M. Hocini, L. Macle, K.-J. Choi, C. Scavee, M. Haissaguerre, and J. Clementy, "Ablation therapy for atrial fibrillation: Past, present, and future," *Cardiovas. Res.*, vol. 54, pp. 337–346, 2002.

[87] G. I. Jee and R. J. Roy, "Adaptive control of multiplexed closed-circuit anesthesia," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 10, pp. 1071–1080, 1992.

[88] I. T. Jolliffe, *Principal Component Analysis*. Springer, 2002.

[89] T. Kaczorek, *Positive 1D and 2D Systems*. London, UK: Springer-Verlag, 2002.

[90] A. S. Kapadia, W. Chan, R. SSSachdeva, L. A. Moye, and L. S. Jefferson, "Predicting duration of stay in pediatric intensive care unit: A Markovian approach," *Eur. J. Oper. Res.*, vol. 124, pp. 353–359, 2000.

[91] E. Kaszkurewicz and A. Bhaya, *Matrix Diagonal Stability in Systems and Computation*. Boston: Birkhaeuser, 2000.

[92] T. Kazama, K. Ikeda, and K. Morita, "The pharmacodynamic interaction between propofol and fentanyl with respect to the suppression of somatic or hemodynamic responses to skin incision, peritoneum incision, and abdominal wall retraction," *Anesthesiology*, vol. 89, no. 4, pp. 894–906, 1998.

[93] T. Kazama, K. Ikeda, K. Morita, M. Kikura, M. Doi, T. Ikeda, and T. Kurita, "Comparison of the effect site KeOs of propofol for blood pressure and EEG bispectral index in elderly and young patients," *Anesthesiology*, vol. 90, pp. 1517–1527, 1999.

[94] F. J. Keefe, D. A. Williams, and S. J. Smith, *Assessment of Pain Behaviors, in* Handbook of Pain Assessment*, 2nd edition*. New York, NY: Guilford: D. C. Turk, and R. Melzack, eds., 2001.

[95] J. Keener and J. Sneyd, *Mathematical Physiology*. New York: Springer-Verlag, 1998.

[96] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.

[97] S. Lankton and A. R. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Imag. Proces.*, vol. 17, pp. 2029–2039, 2008.

[98] P. D. Leenheer and D. Aeyels, "Stabilization of positive linear systems," *Sys. Contr. Lett.*, vol. 44, pp. 259–271, 2001.

[99] J. H. Levy, L. G. Michelsen, J. S. Shanewise, J. M. Bailey, and J. G. Ramsay, "Postoperative cardiovascular management," in *Cardiac Anesthesia* (J. Kaplan, ed.), Philadelphia, PA: WB Saunders, 4 ed., 1999.

[100] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. Available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[101] D. J. C. MacKay, "Bayesian interpolation," *Neural Comp.*, vol. 4, pp. 415–447, 1992.

[102] K. B. Marais and J. H. Saleh, "Beyond its cost, the value of maintenance: An analytical framework for capturing its net present value," *Rel. Eng. Sys. Safety*, vol. 94, pp. 644–657, 2009.

[103] B. Marsh, M. White, N. Morton, and G. N. Kenny, "Pharmacokinetic model driven infusion of propofol in children," *Brit. J. Anaesth.*, vol. 67, no. 1, pp. 41–48, 1991.

[104] L. Martin, *Pulmonary Physiology in Clinical Practice*. St Louis, MO: C. V. Mosby, 1987.

[105] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.

[106] C. J. McGann, E. G. Kholmovski, R. S. Oakes, J. E. Blauer, M. Daccarett, N. M. Segerson, K. J. Airey, N. Akoum, E. N. Fish, T. J. Badger, E. V. R. DiBella, D. Parker, R. S. MacLeod, and N. F. Marrouche, "New magnetic resonance imaging based method to define extent of left atrial wall injury after the ablation of atrial fibrillation," *J. Am. Coll. Cardiol.*, vol. 52, no. 15, pp. 1263–1271, 2008.

[107] S. I. Merkel, J. R. Shayevitz, T. Voepel-Lewis, and S. Malviya, "The FLACC: A behavioral scale for scoring postoperative pain in young children," *Ped. Nurs.*, vol. 23, pp. 293–297, 1997.

[108] K. Murphy, "An introduction to graphical models," 2001, Available from `http://people.cs.ubc.ca/~murphyk/Papers/intro_ gm.pdf`.

[109] K. Murphy, "Bayesian Network Toolbox for MATLAB®," 2002, Available from `http://code.google.com/p/bnt/`.

[110] M. Muzi, R. A. Berens, J. P. Kampine, and T. J. Ebert, "Venodilation contributes to propofol-mediated hypotension in humans," *Aneth. Analg.*, vol. 74, no. 6, pp. 877–883, 1992.

[111] S. A. Nasraway, "The bispectral index: Expanded performance for everyday use in the intensive care unit," *Critical Care Medicine*, vol. 33, no. 3, pp. 685–687, 2002.

[112] S. A. Nasraway, E. C. Wu, R. M. Kelleher, C. M. Yasuda, and A. M. Donnelly, "How reliable in the bispectral index in critically ill patients? a prospective, comparative, single-blinded observer study," *Critical Care Medicine*, vol. 30, no. 7, pp. 1483–1487, 2002.

[113] S. G. Nersesov, W. M. Haddad, and V. Chellaboina, "Optimal fixed-structure control for linear nonnegative dynamical systems," *Int. J. Robust and Nonlinear Control*, vol. 14, pp. 487–511, 2004.

[114] R. G. Newcombe, "Two-sided confidence intervals for the single proportion: Comparison of seven methods," *Stat. Med.*, vol. 17, pp. 857–872, 1998.

[115] M. C. D. Oliveira, J. C. Geromel, and J. Bernussou, "Extended $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norm characterizationsand controller parametrizations for discrete-time systems," *Int. J. Control*, vol. 75, pp. 666–679, 2002.

[116] "OSU SVM Toolbox for MATLAB," `http://sourceforge.net/projects/svm`.

[117] M. Pantic and L. J. M. Rothkrantz, "Toward an affect–sensitive multimodal human-computer interaction," *Proc. IEEE*, vol. 91, pp. 1370–1390, 2003.

[118] J.-F. Payen, O. Bru, J.-L. Bosson, A. Lagrasta, E. Novel, I. Deschaux, P. Lavagne, and C. Jacquot, "Assessing pain in critically ill sedated patients by using a behvioral pain scale," *Crit. Care*, vol. 29, pp. 2258–2263, 2001.

[119] A. Perez, W. Chan, and R. J. Dennis, "Predicting the length of stay of patients admitted for intensive care unit using a first step analysis," *Health Serv. Out. Res. Meth.*, vol. 6, pp. 127–138, 2006.

[120] K. M. Prkachin, "Assessing pain by facial expression: Facial expression as nexus," *Pain Res. Manag.*, vol. 14, pp. 53–58, 2009.

[121] M. A. E. Ramsay and T. M. Savege, "Controlled sedation with alphaxalone–alphadolone," *Brit. Med. J.*, vol. 2, pp. 656–659, 1974.

[122] M. L. Riess, U. A. Graefe, H. V. Aken, and H. G. Bone, "Usefulness of bispectral index to assess the level of sedation in critically ill patients," *Critical Care Medicine*, vol. 27, no. 12, p. A132, 1999.

[123] D. S. Riggs, *Control Theory and Physiological Feedback Mechanisms*. Baltimore, MD: Williams and Wilkins, 1970.

[124] R. G. Ritchie, E. A. Ernst, B. L. Late, J. D. Pearson, and L. C. Sheppard, "Closed-loop control of an anesthesia delivery system: Development and animal testing," *IEEE Trans. Biomed. Eng.*, vol. 34, pp. 437–443, 1987.

[125] T. Sato and K.-Z. Liu, "LMI solution to general $\mathcal{H}_2$ suboptimal controlproblems," *Syst. Control Lett.*, vol. 36, pp. 295–305, 1999.

[126] M. Schiavenato, J. F. Byers, P. Scovanner, J. M. McMahon, Y. Xia, N. Lu, and H. He, "Neonatal pain facial expression: Evaluating the primal face of pain," *Pain*, vol. 138, pp. 460–471, 2008.

[127] T. W. Schnider, C. F. Minto, and D. R. Stanski, "The effect compartment concept in pharmacodynamic modelling," *Anaes. Pharmacol. Rev.*, vol. 2, pp. 204–219, 1994.

[128] P. S. Sebel, E. Lang, I. J. Rampil, P. White, R. C. M. Jopling, N. T. Smith, P. S. Glass, and P. Manberg, "A multicenter study of bispectral electroencephalogram analysis for monitoring anesthetic effect," *Anesth. Analg.*, vol. 84, no. 4, pp. 891–899, 1997.

[129] C. Sessler, M. Gosnell, M. J. Grap, G. Brophy, P. O'Neal, K. Keane, E. Tesoro, and R. Elswick, "The Richmond agitation-sedation scale," *Am. J. Resp. Crit. Care Med.*, vol. 166, pp. 1338–1344, 2002.

[130] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press, 1976.

[131] B. A. Shapiro, "Bispectral index: Better information for sedation in the intensive care unit?," *Critical Care Medicine*, vol. 27, no. 6, pp. 1663–1664, 1999.

[132] L. E. Simmons, R. R. Riker, S. Prato, and G. L. Fraser, "Assessing sedation during intensive care unit mechanical ventilation with the bispectral index and the sedation-agitation scale," *Crit. Care Med.*, vol. 27, pp. 1499–1504, 1999.

[133] J. A. Spain, T. C. Janett, and E. A. Ernst, "The Alabama automated closed-circuit anesthesia project," *Future Anesthesia Delivery Syst.*, vol. 8, pp. 177–183, 1984.

[134] B. Stevens, C. Johnston, P. Petryshen, and A. Taddio, "Premature infant pain profile: Development and initial validation," *Clin. J. Pain*, vol. 12, pp. 13–22, 1996.

[135] A. R. Tannenbaum, *Invariance and System Theory: Algebraic and Geometric Aspects*. Springer-Verlag, 1981.

[136] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, 2000.

[137] SeDuMi 1.1R2, *A Matlab Toolbox for Optimization OverSymmetric Cones*. 2006.

[138] K. Thulasiraman and M. N. S. Swamy, *Graphs: Theory and Algorithms*. Wiley-Interscience, 1992.

[139] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[140] A. W. Toga, *Brain Warping*. San Diego, CA: Academic Press, 1999.

[141] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, 2003.

[142] J. O. Tsokos and C. P. Tsokos, "Statistical modeling of pharmacokinetic systems," *ASM*, vol. 98, pp. 37–43, 1976.

[143] R. N. Upton, G. I. Ludrook, C. Grant, and A. Martinez, "Cardiac output is a determinant of the initial concentration of propofol after short-term administration," *Aneth. Analg.*, vol. 89, no. 3, pp. 545–552, 1999.

[144] V. N. Vapnik, *The Nature of Statistical Learning Theory.* Berlin: Springer–Verlag, 1995.

[145] R. Vidal, Y. Ma, and S. Sastry, *Generalized Principal Component Analysis.* Springer-Verlag, 2005.

[146] R. Vidal, Y. Ma, and S. Sastry, "Generalized principal component analysis (GPCA)," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, pp. 1945–1959, 2005.

[147] R. Vishnoi and R. J. Roy, "Adaptive control of closed-circuit anesthesia," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 1, pp. 39–47, 1991.

[148] K. Q. Weinberger and L. K. Saul, "Unspurvised learning of image manifolds by semi-definite programing," *Int. J. Comp. Vis.*, vol. 70, pp. 77–90, 2006.

[149] A. A. Weinbroum, R. B.Abraham, T. Ezri, and J. Zomer, "Wrist actigraphy in anesthesia," *Journal of Clinical Anesthesia*, vol. 13, pp. 455–460, 2001.

[150] D. R. Westenkow and W. S. Jordan, "The Utah system: Computer controlled anesthesia delivery," *Future Anesthesia Delivery Syst.*, vol. 8, pp. 221–233, 1984.

[151] R. T. Whitaker, "A level-set approach to 3D reconstruction from range data," *Int. J. Comp. Vis.*, vol. 29, no. 3, pp. 203–231, 1998.

[152] A. Y. Yang, S. Rao, A. Wagner, Y. Ma, and R. M. Fossum, "Hilbert functions and applications to the estimation of subspace arrangements," Int. Conf. Comp. Vis., Beijing, China, pp. 158–165, 2005.

[153] A. Yezzi, A. Tsai, and A. Wilsky, "A statistical approach to snakes for bimodal and trimodal imagery," in *Proc. Int. Conf. Comp. Vis.*, pp. 898–903, 1999.

[154] L. A. Zadeh, "The role of fuzzy logic in management of uncertainty in expert systems," *Fuzzy Sets Syst.*, vol. 11, pp. 199–227, 1983.

Vita

Behnood Gholami received the B.Sc. and M.A.Sc. degrees in mechanical engineering from the University of Tehran, Tehran, Iran, and Concordia University, Montreal, Canada, in 2003 and 2005, respectively. In 2009 he received the M.S. degrees in mathematics and aerospace engineering from the Georgia Institute of Technology, Atlanta, GA. Since 2005 he has been with the the School of Aerospace Engineering, Georgia Institute of Technology, where he is currently working toward the Ph.D. degree in aerospace engineering. In 2007 he served as a summer intern with the Control & Identification Group, The MathWorks, Inc., Natick, MA. He was a recipient of the Concordia University Graduate Fellowship and a Power Corporation of Canada Graduate Fellowship.

His research interests include computer vision, machine learning, vision-based control, medical image processing, optimal control, robust control, and model predictive control. Applied areas of interest include biomedical and biological systems, active control for clinical pharmacology, and image-guided therapy.